Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 2, 372–385
https://doi.org/10.1037/xge0001491

The Construction and Use of Cognitive Maps in Model-Based Control

Ata B. Karagoz, Zachariah M. Reagh, and Wouter Kool
Department of Psychological and Brain Sciences, Washington University in St. Louis

When making decisions, we sometimes rely on habit and at other times plan toward goals. Planning requires the construction and use of an internal representation of the environment, a cognitive map. How are these maps constructed, and how do they guide goal-directed decisions? We coupled a sequential decision-making task with a behavioral representational similarity analysis approach to examine how relationships between choice options change when people build a cognitive map of the task structure. We found that participants who encoded stronger higher-order relationships among choice options showed increased planning and better performance. These higher-order relationships were more strongly encoded among objects encountered in high-reward contexts, indicating a role for motivation during cognitive map construction. In contrast, lower-order relationships such as simple visual co-occurrence of objects did not predict goal-directed planning. These results show that the construction of cognitive maps is an active process, with motivation dictating the degree to which higher-order relationships are encoded and used for planning.

Public Significance Statement

This study combines behavioral measures and computational modeling to show that people build internal representations of a task—cognitive maps—and that the quality of their maps predicts how much they plan toward reward. We found that these maps were abstract, encoding relationships beyond those derived from immediate experience, and that they were more accurate when more reward was at stake. Our findings show that people's ability and motivation to plan toward goals relate to their construction and use of cognitive maps.

Keywords: cognitive map, behavioral representational similarity, reinforcement learning, two-step task

Supplemental materials: https://doi.org/10.1037/xge0001491.supp

A fundamental question of human cognition is how we are able to effectively plan toward goals. Decisions are not made in a vacuum, but rather capitalize on the structure of the world around us. Imagine moving to a new city and learning about its structure as you navigate it. Ideally, your internal model will be set up to guide effective planning, strongly encoding relationships among major streets that connect different neighborhoods, especially when you anticipate visiting them frequently. However, this model may also incorporate features that are less relevant for planning. For example, you may encode relationships between streets with similar-sounding names or that happen to be located close together. Here, we ask how humans learn which features of the environment are relevant for planning and apply these representations to optimize behavior. In other words: how do we build and use internal models to guide our decisions?

These internal models are commonly referred to as "cognitive maps," a term coined by Tolman to explain how rodents were able to use spatial features of a maze to navigate toward goals (Tolman, 1948). Since Tolman, specific cells that represent both current (Hafting et al., 2005; O'Keefe & Nadel, 1978) and possible future

This article was published Online First December 7, 2023.
Ata B. Karagoz https://orcid.org/0000-0002-7032-0355
We would like to thank Michael Freund, Gabriel James, Jeff Zacks, and Bettina Bustos for helpful conversations. We also would like to thank the members of the Complex Memory Lab and the Control and Decision Making Lab for all their support. Our work was supported in part by the McDonnell Center for Systems Neuroscience. The content is solely the responsibility of the authors and does not necessarily represent the official views of our supporting organizations. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the article. The authors have no conflicts of interest to declare. The data in this article were presented as a poster at the 2022 Reinforcement Learning and Decision Making (RLDM) conference in Providence, Rhode Island. Wouter Kool presented this work at the TU Dresden Symposium on "Volition and Cognitive Control: From Executive Functions to Meta-Control." Zachariah M. Reagh presented this work at the annual meeting of the Memory Disorders Research Society (MDRS) as well as at the Park City Winter Conference on Learning and Memory. Most of the text is different, the figures have changed, and the added data and analyses have not been disseminated anywhere.

Ata B. Karagoz served as lead for conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, validation, visualization, writing–original draft, and writing–review and editing. Zachariah M. Reagh and Wouter Kool contributed equally to resources, writing–review and editing, and supervision.

Correspondence concerning this article should be addressed to Ata B. Karagoz, Department of Psychological and Brain Sciences, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, United States. Email: a.b.karagoz@wustl.edu

spatial locations (Johnson & Redish, 2007) have been discovered, providing the neural underpinnings for cognitive maps. Critically, recent work has applied the concept of cognitive maps beyond spatial domains and toward nonspatial and perhaps even very abstract relationships (see Behrens et al., 2018 for review). In this view, cognitive maps can represent social relationships (Park et al., 2021; Tavares et al., 2015), transitions between tasks (Schuck & Niv, 2019; Schuck et al., 2016), and abstract associations between individual items in a task (Constantinescu et al., 2016; Theves et al., 2020). However, much remains unclear about how humans construct abstract cognitive maps, how they use them to plan toward goals, and whether the quality of these maps relates to goal-directed behavior.

Recent studies of decision-making formalize goal-directed planning over a cognitive map as "model-based" reinforcement learning (RL; Daw et al., 2005; Drummond & Niv, 2020; Gläscher et al., 2010; Sutton & Barto, 2018). A model-based learner computes expected values for available actions by using a representation of the task structure. This flexible but computationally costly strategy enables it to apply the values learned at a given goal to every path that leads to that goal. Meanwhile, a model-free learner uses a more efficient but inflexible strategy, only updating the value of actions that led to reward, without considering the structure of the task.

Here, we assess the construction of cognitive maps as a driving force of goal-directed decisions. To do so, we had participants perform a variant of the "two-step" task (Daw et al., 2005, 2011; Kool et al., 2016, 2017). This task dissociates model-based and model-free control by exploiting the ability of the model-based system to plan using an internal representation of the task structure, which contrasts with the model-free reliance on direct action–reward associations. Many prior studies using this task assume or ensure that an effective representation of the task is present (but see Feher da Silva & Hare, 2020; Feher da Silva et al., 2023). Here we specifically trained our participants only briefly on the rules of the task and not the actual transition structure to try and track individual differences in cognitive map formation. These individual differences in cognitive maps may critically influence differences in model-based control. How does the nature and quality of one's cognitive map influence behavior?

To index cognitive map structure in a purely behavioral setting, we developed a novel approach, which we term behavioral representational similarity analysis (behRSA). Our approach was inspired by neuroimaging analyses (Dimsdale-Zucker & Ranganath, 2018; Kriegeskorte et al., 2008) that assess pairwise neural pattern similarity between stimuli. In our case, participants simply rated the perceived relatedness of pairs of items before and after learning the two-step task. This allowed direct insights into representational similarity among these components as defined by the participants' own behavioral output. We assessed the amount of planning-relevant versus planning-irrelevant information incorporated into their maps, and whether motivation influenced map formation.

We found that participants whose similarity ratings reflected the higher-order transition structure earned more reward and used more model-based control, but we found no such relationship for components reflecting more superficial structure. Moreover, such structure learning was amplified for contexts in which more reward was at stake. This study introduces a method for inferring internal representations on the basis of behavioral ratings alone and provides insight into the way people construct and plan over these representations.

Method

Preregistration

This article describes a preregistered replication (https://osf.io/uw3p7) of a pilot study. This pilot sample can be seen in the online supplemental materials, and the data are displayed in Figures S1 and S2 in the online supplemental materials.

Participants

We recruited 209 healthy younger (range = 18–35 years old) adults for this study using the Cloud Research Platform and Amazon Mechanical Turk. We excluded 48 participants based on preregistered criteria. We excluded 32 participants because they failed to respond in more than 20% of the trials in the decision-making task. We excluded 12 participants because they failed to respond in more than 20% of trials in the similarity task. We also excluded participants if they engaged in purely random action selection (Patil et al., 2021). To determine this, we simulated 1,000 agents that responded randomly on each trial, yielding a distribution of log-likelihoods expected under random action selection. We excluded two participants whose best-fitting log-likelihood exceeded the fifth percentile of this distribution (negative log-likelihood score $>$ 172.846). We also excluded two participants because of missing data in their behavioral representational similarity task, which rendered certain pairwise comparisons impossible. This resulted in an effective sample of 161 younger adults (102 male, 58 female, and one nonbinary, age range = 18–35 years, $M_{\text{age}} = 29.2$ years). We stopped collecting data once our effective sample size reached 160, which allowed us to detect a correlation with the same strength of our initial study ($r = .2764$) with 95% power.

All participants were compensated with a base payment of $9 and additional performance-related payments of 1 cent for every 25 points obtained in the decision-making task. All participants gave informed consent, and procedures were approved by the Washington University in St. Louis Institutional Review Board.

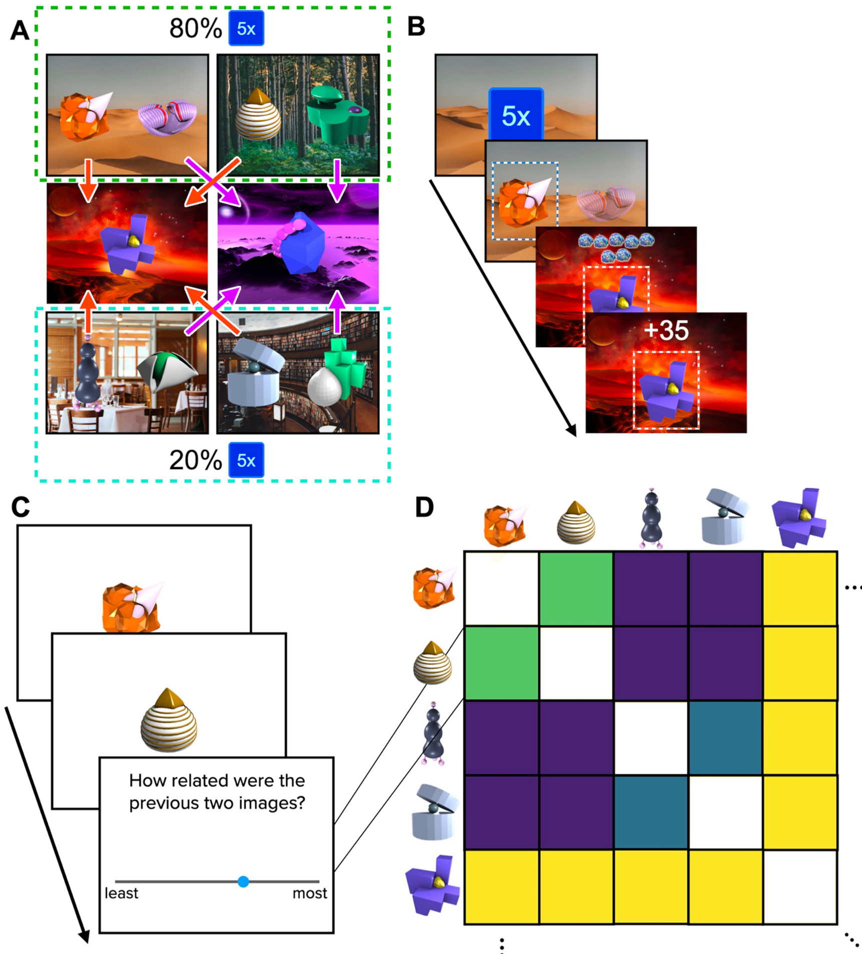
Behavioral Representational Similarity Task

In order to test how task experience and motivation affect structure learning, participants provided relatedness ratings of pairs of novel 3D objects that were also used as choice options in the decision-making task (adapted with permission from Hsu et al., 2014; Schlichting et al., 2015). On each trial, they were shown an initial object for 1 s, followed by a 500 ms fixation cross, and then another object for 1 s. Participants were then shown a "slider" bar and were asked to move the slider to their perceived level of "relatedness" between the two objects they had just seen (Figure 1C). Participants had 5.5 s to provide a response. All objects were presented in the center of the screen, with image sizes of 400 × 400 pixels. Participants performed all pairwise ratings of all 10 objects in both orders, resulting in 90 trials. Participants performed the behavioral representational similarity task both before and after the decision-making task. This allowed for a pre- versus postlearning comparison of object-relatedness judgments.

Decision-Making Task

The decision-making task was designed to dissociate between model-free and model-based decisions in a setting where

Figure 1
Task Schematic



Note. (A) Task transition structure. Four distinct first-stage states (depicted in the top and bottom row) each contain two unique choice objects. Each of these objects deterministically leads to one of two second-stage states (depicted in the middle row), as depicted by the colored arrows. These first-stage states were associated with different amounts of reward that changed across the task. For the two "high-reward" states (top row), 80% of the trials involved a high stake, with a multiplier cue indicating that any points would be multiplied by five. For the other two "low-reward" first-stage states (bottom row), high-stake trials occurred on only 20% of trials. (B) Representative trial. First, the participant is shown the first-stage state and the stake multiplier and then is shown the objects. After selecting one of these, the trial transitions to the corresponding second-stage state, where they receive seven space treasure pieces after selecting the second-stage object. The high-stake multiplier amplifies this to 35 points. (C) Representative trials of the behRSA task. Participants are shown each novel object pairing twice. (D) A portion of a hypothetical matrix of similarity ratings generated from the behRSA task. behRSA = behavioral representational similarity analysis. See the online article for the color version of this figure.

participants need to learn about the structure of the task, and was based on a recently developed "two-stage" sequential decision-making task (Kool et al., 2017). Each trial of the task started randomly in one of four first-stage states. Each of these states offered a choice between a unique pair of "teleporters," presented side-by-side. Participants used the "F" key on their keyboard to choose the left teleporter, and the "J" key to choose the right teleporter. This choice determined which one of two second-stage states would be encountered. For each pair, one of the teleporters deterministically led to a purple second-stage state, and the other deterministically led to a red second-stage state. Importantly, each teleporter always led to the same second-stage state (Figure 1A).

Each second-stage state contained a unique "generator" that was associated with a scalar reward. Participants were instructed to press the spacebar key to interact with the generator so that it provided them with "space treasure," and they were told that the fuel rods used by the generators would sometimes yield more or less space treasure. The payoffs of the generators changed over the course of the experiment according to independent random walks. Their reward distributions were initialized randomly for each participant within a range of 0–9 points and then varied according to a Gaussian random walk ($\sigma = 2$) with reflecting bounds at 0 and 9.

This task distinguishes between model-based and model-free strategies, since only a model-based decision-maker generalizes experiences from one starting state to all other starting states. That is, after receiving a high reward in a second-stage state, a model-based learner can use its knowledge of the transitions to plan its way back to that same second-stage state. A model-free agent, on the other hand, learns through action–reward associations and will only become more likely to choose that same action in the same first-stage state, not transferring experiences from one first-stage state to the others (Kool et al., 2016). The computational model, described below, uses RL to capture this distinction in a single model-based weighting parameter. The 10 objects were randomly assigned as teleporters and generators for each participant separately.

The first-stage states were not only uniquely identifiable by the pair of teleporters, but they also were associated with a set of background images. Specifically, each of the first-stage states was always presented with a background image that belonged to one of four "location" categories. Two of the first-stage states were always shown with background images of "inside" locations (libraries and restaurants), and the other two with images of "outside locations" (deserts and forests). Each of these categories contained four images. Their presentation was selected pseudorandomly, so that each background image was shown equally frequently.

In order to introduce differing incentives for model-based control and demands for learning the task structure, we introduced a "stakes" manipulation in this task (Kool et al., 2017). At the start of each trial, an incentive stake cue indicated by how much the reward obtained at the end of the trial would be multiplied. On some trials, this cue indicated that the points would be multiplied by 5 (*high stakes*). On other trials, the cue indicated that the points would be multiplied by 1 (*low stakes*). For example, if a participant earned five space treasure pieces on a high-stakes trial, the multiplier would result in a total of 25 points. On a low-stakes trial with the amount of space treasure, the participant would earn five points. Importantly, the chances of a high-stake trial were different between the first-stage states. For each participant, either the "inside" or "outside" first-stage states were selected as "high-stakes" states, and the other as the "low-stakes" states. In the high-stakes states (high-stake context), 80% of trials had a 5× multiplier and 20% of trials had a 1× multiplier. In the low-stakes states (low-stake context), 20% of trials had a 5× multiplier and 80% of trials had a 1× multiplier.

At the start of each trial, participants saw the category background and the stake multiplier for 1 s. Then, the stake moved to the top left corner and the teleporters were presented, and participants were given a time limit of 1.5 s to choose between them (following prior work by Kool et al., 2017). After the response, the selected option was highlighted, and the nonselected option was grayed out for the remainder of the response period. There was a 500 ms interval between the end of the first-stage response period and the onset of the second stage. Following a 200 ms interval after the generator was selected, participants were shown how many space treasure pieces they earned for 1.5 s. Then each piece was converted to points and added to the total score (100 ms each point). There was a 500 ms intertrial interval (ITI). Participants completed a total of 256 trials with an optional short break in the middle.

Before performing the decision-making task, participants were briefly instructed on its rules and were given practice, but not the specific transition structure. They were first familiarized with two first-stage practice contexts and two second-stage practice contexts. Each of these first-stage contexts was associated with a pair of unique practice teleporter objects and each of the second-stage contexts was associated with a single unique space generator object. Participants learned that in each first-stage state, the teleporter objects appeared side-by-side and that each led deterministically to one of the two second-stage planets. Participants then learned which object led to which second-stage state through practice. We informed participants that the practice objects and contexts were different from those used in the main task, but that the rules governing the transition structure would be similar. It is important to note that the participants did not experience any of the stimuli, or the task structure, of the main phase for the decision-making task and, that they were only able to learn it during task performance.

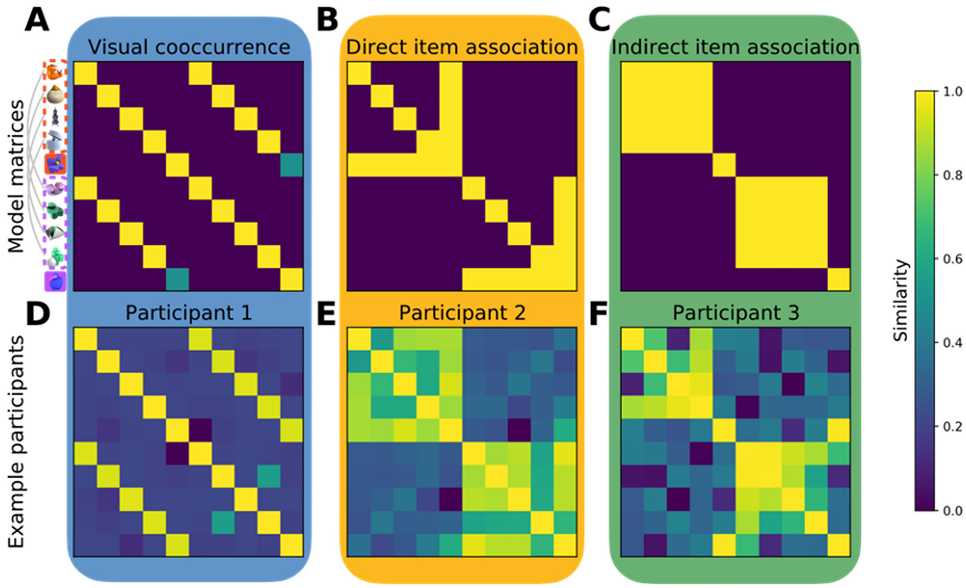
Memory Task

Participants also completed a surprise memory test where they were shown a pairing of a background image and a pair of objects. They were asked to indicate whether they saw this pairing before. Specifically, they were given 5 s to respond on a 4-point scale consisting of "*yes, sure*," "*yes, unsure*," "*no*, *unsure*," and "*no, sure*." After a 500 ms ITI, a new stimulus/background pairing was shown. Participants were shown all 16 background/object pairings they saw in the task (4 for Each Planet × 4 Planets). They were also presented with "lure" trials designed to test resistance to memory interference. In 12 of these, participants saw choice objects of one state presented on a background of another state (i.e., the desert objects in front of a forest they have seen), which we termed mismatch trials. In the remaining 16 lure trials, a pair of choice objects were shown on a new image from each category of background (i.e., desert objects in front of a desert they have never seen), which we termed lure trials. This resulted in a total of 44 trials.

behRSA

We used participants' relatedness ratings of objects to measure the structure of their cognitive maps (Figure 1C and D). Specifically, we hypothesized that experience with the decision-making task would yield participants to represent the task structure across three different levels of abstraction. To test this hypothesis, we compared their matrices of relatedness ratings to three a priori model matrices, which reflected these levels of representation of the decision-making task structure (Figure 2A–C). First, we hypothesized that objects could become related if they co-occurred in a first-stage state (e.g., the two teleporters from the desert first-stage state; Figure 2A). Even though this representational shift, which we call "visual co-occurrence," captures some task structure, it is not useful for planning since it is unrelated to the consequences of actions. There is also increased similarity for the two items in the second-stage states, because we reasoned people would infer they share a "reward state." We also hypothesized two forms of representational shifts that related to the task's transitions (Figure 2B and C). First, we hypothesized that first-stage state objects and the second-stage

Figure 2
Model Matrices and Example Participants



Note. Representational similarity matrices. The top row depicts hypothetical model matrices. The bottom row depicts sample behRSA data from participants that are well captured by the corresponding model matrix. In each matrix, each cell corresponds to a similarity measure between two objects (defined by row and column). The diagonal cells correspond to self-similarity and are assigned full similarity by default. (A) Visual co-occurrence model of representational change in which objects that appear together in the same first-stage state become highly similar, and objects that appear in second-stage states become mildly similar. (B) Direct item association model of representational change where first-stage objects become more similar to the second-stage item they lead to. (C) Indirect item association model where first-stage objects that all lead to the same second-stage state become more similar to each other. (D) Example participant with high visual co-occurrence model fit. (E) Example participant with high direct item association model fit. (F) Example participant with high indirect item association model fit. behRSA = behavioral representational similarity analysis. See the online article for the color version of this figure.

objects they lead to would become related (i.e., the teleporter from the desert that led to the red planet and the red planet's generator). This representation, which we call a "direct item association," allows for planning from a first-stage action to the related second-stage state. Finally, we hypothesized that all first-stage objects leading to the same second-stage state would become related. We call this representation an "indirect item association," because it encodes relations between objects that never occurred on the same trial. Critically, these indirect associations would indicate that participants are abstracting beyond immediate experience in a way that goes above a simple understanding of action-outcome contingencies in the task. Note that these model matrices differ in the degree to which they reflect associations that are useful for goal-directed control. Specifically, the direct outcome association and the indirect item-outcome models encode the consequences of choices, whereas the visual co-occurrence model encodes a more superficial aspect of the task that does not reflect such higher-order structure.

For each model matrix, cells that reflect the hypothesized relatedness relationship were coded as 1, and cells that did not were coded as 0 (with the exception that the two second-state stages were coded as 0.5 with relation to one another in the visual co-occurrence model, due to shared "reward state"). It is important to note that these model matrices are agnostic to the identity of specific objects in the task structure, because these were randomly assigned for each participant. In sum, each matrix cell indicates a pairwise similarity rating between two objects, and the hypothesized matrices reflect three predicted ways in which participants could change their perceived similarity among objects based on learning this task structure.

We calculated the mean relatedness ratings for each pair of objects separately for the pre- and posttask ratings. We then subtracted the pretask means from the posttask means to get a measure of representational change. Finally, we fit a multiple regression model for each participant using the three hypothesized models of behavioral similarity, allowing us to extract β coefficient values for each participant's data fit to each model matrix:

$$\begin{aligned}\text{Participant Similarity Matrix } = \; & \bar{\beta}_{\text{visual co-occurrence}} \times \text{Visual Co-Occurrence} \\ & \text{Model Matrix} + \bar{\beta}_{\text{direct item}} \times \text{Direct} \\ & \text{Item Model Matrix} + \bar{\beta}_{\text{indirect item}} \times \\ & \text{Indirect Item Model Matrix}\end{aligned} \tag{1}$$

These three regression coefficients each reflect the strengths of one of the hypothesized representational shifts.

We also tested whether the task structure was learned differently for the high-stakes and low-stakes first-stage states. To do so, we split the model relatedness matrices along high-stakes contexts and low-stakes contexts and fit separate multiple regression models, resulting in six coefficient values per subject (one for the low-stakes and high-stakes arm for each of the three models). We then subtracted the high-stakes arm coefficient from the low-stakes arm coefficient to provide a difference score (Figure 3B).

RL Model

We used a dual-systems RL model to examine participants' use of model-based control (see the online supplemental materials). This model combines model-free and model-based learning systems that learn state-action values. The model-free system uses a temporal difference-learning algorithm to heighten values for actions in response to positive prediction error and lower values that lead to negative prediction errors. The model-based learner combines the transition structure of the task with second-stage model-free values to plan its first-stage choices. These two learning systems are combined using a weighting parameter (w) bounded between 0 and 1, where 0 is fully model-free control and 1 is fully model-based. The combined system then made choices according to an inverse-temperature parameter β that governed the explore exploit tradeoff between the two first-stage options, where a value of 0 dictates pure exploitation and higher values result in more exploration and value insensitivity. The model also included a learning-rate α that governed updates to stored values after rewards, an eligibility trace parameter λ that controls how the outcome at the second stage informs the first-stage, parameters π and ρ which capture perseveration on either response or stimulus choice, and finally η and κ which determine transition-matrix updates.

We used maximum a posteriori estimation to fit this dual-system RL model to behavior on this task (see online supplemental materials for more details). The estimation method used empirical priors previously reported by Bolenz et al. (2019).

We would like to note that the mixture model simply describes the relative strength of the two systems across trials, and should not be interpreted as a strict description of trial-level behavior. When participants use model-based control on a larger proportion of trials, their fitted mixture parameter w increases.

Exploratory Principal Components Analysis (PCA)

In addition to the a priori model matrices described above, we were also interested in using a more data-driven approach to capture structure in relatedness ratings. Therefore, we used a PCA to reconstruct the data-driven model matrices that would describe the most shared variance among the participants. To do this, we first flattened into a vector the entries of participants' representational similarity matrix that are below the main diagonal. These vectors were combined into a Subject × Data Matrix which we extracted the principal components using singular value decomposition. This plot indicates which of the principal components explain more than 5% of the variance, which were the first three principal components in our data. We then took each of these components and remapped them into the original symmetric matrix, and visually inspected them to get an assessment of what the representational change along that principal component dimension (Figure 4A).

Exploratory Multidimensional Scaling Analysis

As a complement to the exploratory PCA, we also performed a multidimensional scaling analysis on the change from pre- to postexposure representational similarity ratings for each participant. To do this, we first computed matrices of the change in similarity ratings from pre- to postexposure for participants. Then, after removing the diagonal, we computed the average change for each item across all participants, resulting in a single 10 × 10 similarity matrix that is symmetrical around the diagonal. Next, we submitted this similarity matrix into MATLAB's mdscale.m function, requesting a multidimensional scaling solution over two dimensions. In short, this technique provides co-ordinates in a 2D space for each item, with Euclidean distances between items revealing subjective task structure. We then visually inspected these solutions for relationships among items and with the reward context (Figure 4C).

Analysis of Decision-Task Data

We computed average performance on the decision-making task as the average number of points earned per trial. To correct for baseline differences in available reward (as a result of the random Gaussian walks), we then subtracted the average available reward across both second-stage states. Participants' data were also fit using a RL model (described in RL model).

Analysis of Memory Probe Data

We used the sensitivity index d' to assess participant's memory precision during the memory probe task. To compute this, the first one obtains the hit rate (H) which is the proportion of target trials that a subject correctly identifies as old. Next one calculates the false alarm rate (F) which is the proportion of lure trials that a subject incorrectly identifies as old. Then, the d' measure is calculated as the following equation, $d' = z(H) - z(F)$, where $z(X)$ is the z score. We used this equation to calculate two separate d' measures, one for lure trials and one for mismatch trials. The first was calculated using a false alarm rate for true lure trials, which we call d'_{lure}. The second was calculated using a false alarm rate from the mismatch trials, we called this d'_{mismatch}. We calculated these measures separately for high- and low-stakes contexts. The d'_{mismatch} value was then compared for the high- and low-stakes contexts using a paired-sample t test.

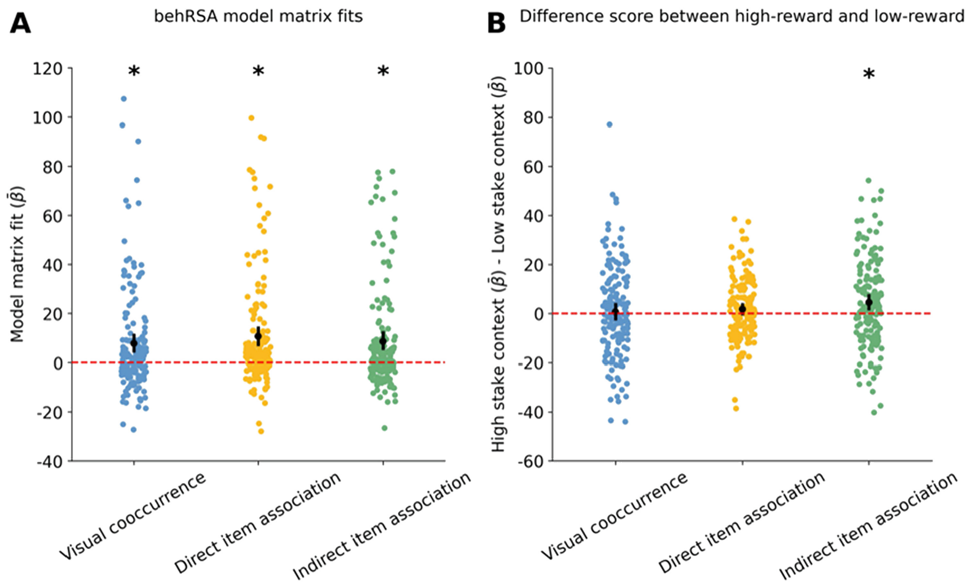
Transparency and Openness

The data for this study are publicly accessible (at https://www.github.com/cdml-lab/mb-cog-maps-paper). The code is also publicly accessible (at the same link). Most of the materials are publicly available here (https://osf.io/hr4qk/), however, we are unable to share the novel object stimuli since we did not get permission from the original creators. Our study's design, hypotheses, and analysis plan were preregistered here (https://osf.io/uw3p7).

Results

We combined a RL task with behRSA to measure how experience and motivation influence the representation of task structure. Participants ($n = 161$) performed pairwise similarity ratings on 10 novel objects before and after encountering them in a sequential decision-making task that distinguishes model-free and model-based control (Figure 1). Importantly, we used a variant of this task that is designed such that increased use of model-based control is linked to increased performance (Kool et al., 2016). This is a

Figure 3
Model Matrix Fits



Note. (A) Participants' model matrix fits for the three hypothesized models. Model matrix fit is calculated using a multiple regression that produces three coefficients for each participant, one for each hypothesized model. All three of the models are represented within the sample. (B) Participants on average show a higher degree of indirect item association for the high-stake versus low-stake context. * represents $p < .05$, error bars are 95% CI. behRSA = behavioral representational similarity analysis; CI = confidence interval. See the online article for the color version of this figure.

preregistered replication of a prior experiment (see Methods: Cohort 1 in the online supplemental materials).

Model-Fitting Results

Mirroring prior work (Daw et al., 2011; Kool et al., 2017), we found that participants' behavior reflected a mixture of model-free and model-based control (mean $w = 0.57$). This suggests that participants learned an internal representation of the task, a cognitive map, and that they used this for goal-directed decision-making.

Behavioral Representational Similarity Analysis (RSA)

At the group level, participants judged item similarity in a manner consistent with each of the three model matrices (example subjects with strong fits to each coefficient can be seen in Figure 2D–F, respectively). As depicted in Figure 3A, participants judged objects as more related when they had occurred in the same first-stage state, $t(160) = 4.64$, $p < .001$, $d = 0.37$, when they constituted a pair where one object transitioned to the other, $t(160) = 5.98$, $p < .001$, $d = 0.47$, and when they both led to the same second-stage state, $t(160) = 5.29$, $p < .001$, $d = 0.42$. In other words, all three hypothesized components of the task were represented at the group level.

Exploratory Data-Driven Analyses

Next, we conducted a data-driven test of whether our hypothesized components captured the primary patterns that emerged from the relatedness ratings. We performed a PCA on the aggregate behRSA data and inspected the three dimensions that described the most variance (Figure 4A). The results were largely consistent with our hypothesized relationships. The first principal component corresponded to a mixture of our models of direct and indirect item association. The second corresponded to a combination of dissimilarity of co-occurrence, coupled with increased similarity for the direct and indirect item associations. Finally, the third principal component resembled a combination of increased similarity for visual co-occurrence and direct item associations. This provides strong support for our predicted representations of task structure.

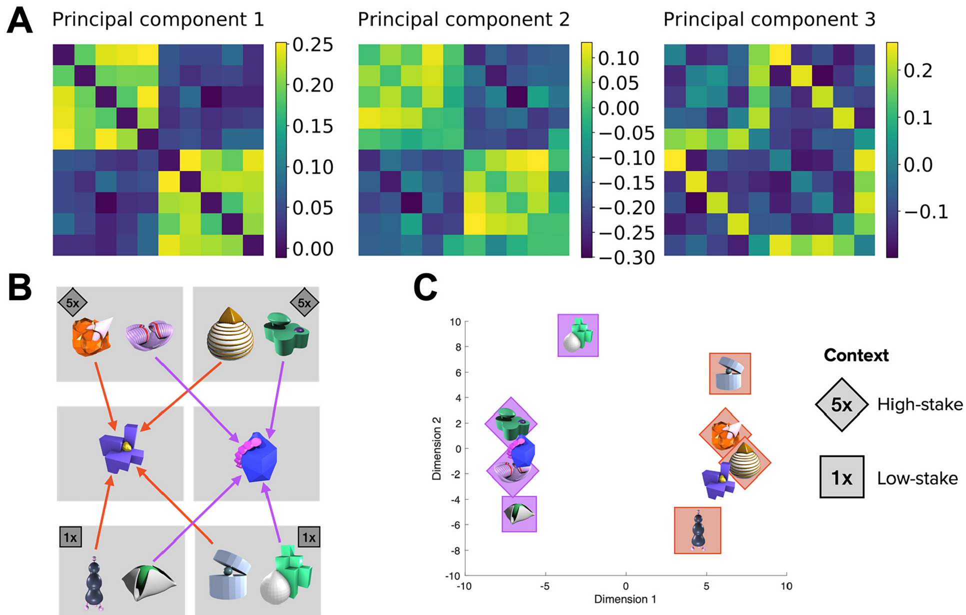
We also performed a multidimensional scaling analysis (Figure 4C), so as to assess the relationship between items in a reduced, 2D, similarity space. This analysis revealed several interesting relationships. First, we observed that items that did not share a relationship with a goal state were separated along the primary dimension. Second, we observed that the reward context drove representational distance along the secondary dimensions, with stimuli from high-reward contexts fitting closer to the goal-state than stimuli from low-reward contexts. These findings mirror not only the primary behRSA measures reported above, but also their interaction with reward-context described below.

Representations of Task Structure Correlate With Task Performance and Model-Based Control

If the behRSA results reflect different aspects of a cognitive map, and cognitive maps enable planning, then individual

Figure 4

Exploratory Analysis of Similarity Ratings



Note. (A) Inspecting the first three principal components of participants' behRSA data we find that the principal component that drives the most variance (Principal component 1) is similar to an equal mixture of our a priori direct and indirect item association models. The second principal component reflects a mixture of a negative similarity for object's visual co-occurrence combined with indirect and direct item associations. Finally, the third principal component is a mixture of direct item association and visual co-occurrence models. Color bar depicts similarity structure in arbitrary units. (B) Simplified task figure representation. (C) Multidimensional scaling visualization of overarching principles in participant behRSA. The teleporters from the high-stakes contexts are clustered close to their respective second-stage generators. The teleporters from the low-stakes contexts are further out along Dimension 2. Axes are arbitrary units. behRSA = behavioral representational similarity analysis. See the online article for the color version of this figure.

differences in these representations of similarity structure should correlate with task performance (average reward rate) and reliance on model-based control (Figure 5). We predicted positive correlations for aspects of the task structure that are important for goal-directed planning, but not for lower-order relationships.

Consistent with our hypotheses, we did not observe a relationship between performance (i.e., points earned in the task) and the strength of the visual co-occurrence component, $r(159) = .004$, 95% CI [−0.15, 0.16], $p = .96$. However, we found that performance was positively correlated with the encoding of direct item associations, $r(159) = .32$, 95% CI [0.18, 0.46], $p < .001$ and indirect item associations, $r(159) = .45$, 95% CI [0.31, 0.56], $p < .001$.

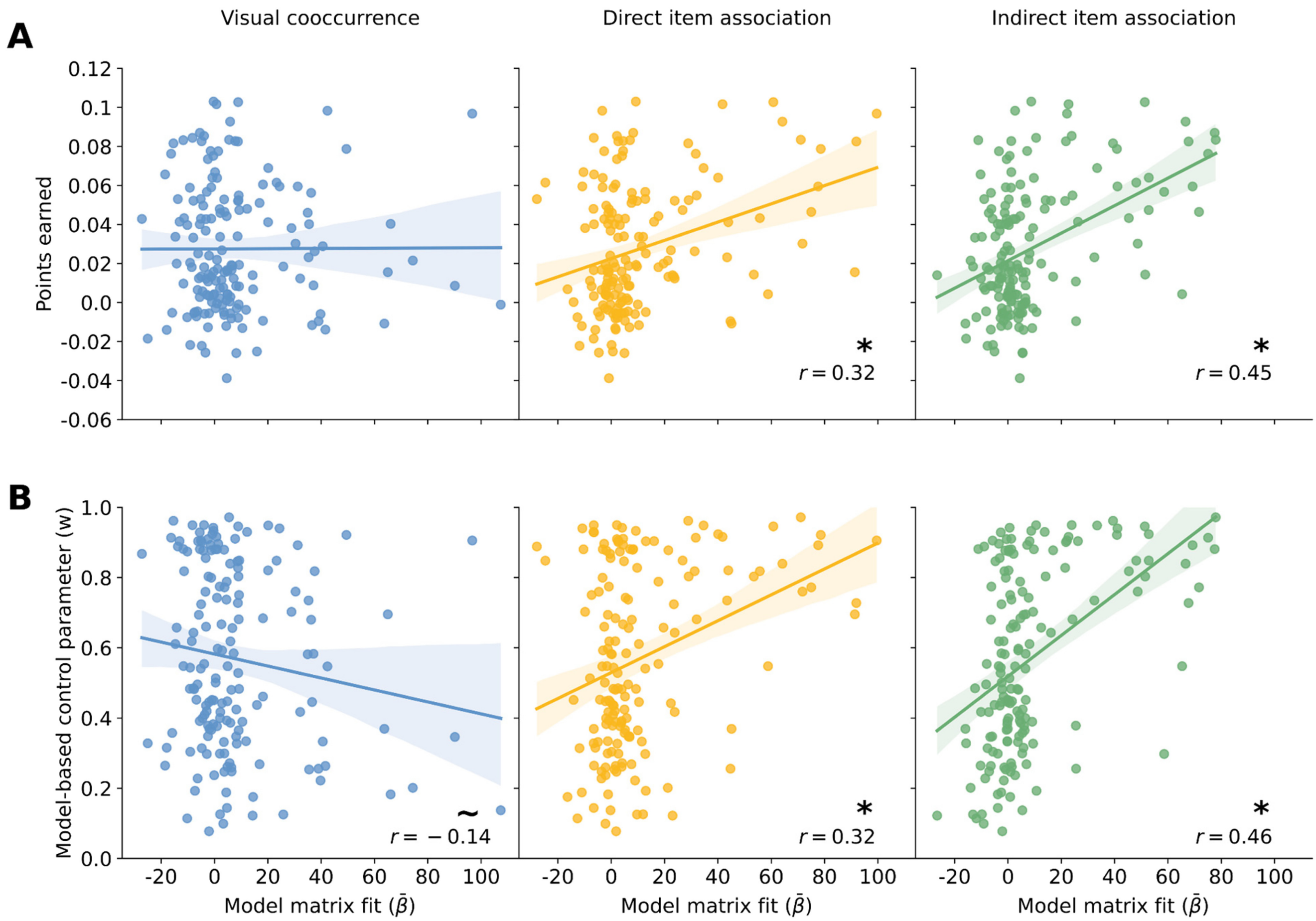
Next, we found a trending but nonsignificant negative correlation between model-based control and the strength of the visual co-occurrence component, $r(159) = -.14$, 95% CI [−0.29, 0.01], $p = .0753$. Critically, however, we found that the use of model-based control was positively correlated with the encoding of direct item associations, $r(159) = .33$, 95% CI [0.18, 0.46], $p < .001$, and indirect item associations, $r(159) = .48$, 95% CI [0.35, 0.59], $p < .001$.

In order to assess the difference in strength between these correlations, we ran a series of Williams' tests to compare them. First, we compared the correlations of the model matrix fits with the performance measure (points earned) The correlation between points earned and visual co-occurrence was weaker than its correlation with the indirect item association, Williams' $t(158) = -4.39$, $p < .001$, and direct item association, Williams' $t(158) = -3.67$, $p < .001$. Also, its correlation with indirect item association was weaker than its correlation with the direct item association, Williams' $t(158) = 2.10$, $p = .036$.

We also performed these tests for the correlations between model-based mixture weight w and the behRSA parameters. Its correlation with visual co-occurrence was weaker than its correlation with the indirect item association, Williams' $t(158) = -2.4670$, $p < .001$, and the direct item association, Williams' $t(158) = -5.5955$, $p < .001$. Again, its correlation with the indirect item association was weaker than its correlation with the direct item association, Williams' $t(158) = 2.4670$, $p = .0143$. Thus, for both points earned and estimates of model-based control, we found the strongest link to the most abstracted components of the cognitive map.

Inspired by a reviewer's comments, we also performed a series of multiple linear regressions to account for shared variance among the different portions of participants' cognitive maps. First, a linear regression was conducted to examine the relationship between performance in the decision-making task and participants' behRSA matrix components. The model included 161 observations and had an R^2 of .203 and an adjusted R^2 of .188. The results indicated a

Figure 5
Model-Based Representations Correlate With Decision-Making Task and RL Model



Note. (A) Comparison of behRSA representations of subjects with their measure of performance in the decision-making task. Direct item associations and indirect item associations correlate with performance in the decision-making task whereas simple visual co-occurrence does not. (B) Reconstruction of task-relevant representations in behRSA also correlates with increased use of model-based control, whereas visual co-occurrence does not[1]. RL = reinforcement learning; behRSA = behavioral representational similarity analysis. See the online article for the color version of this figure.
* $p < .05$.

significant positive relationship between performance in the decision-making task and the indirect item association component of participants' behRSA, $b = 0.013$, $SE = 0.003$, $t(157) = 4.118$, $p < .001$. There were no significant relationships between task performance and direct item association, $b = 0.003$, $SE = 0.003$, $t(157) = 0.879$, $p = .381$. Visual co-occurrence also had no significant relationship with task performance, $b = -0.001$, $SE = 0.003$, $t(157) = -0.287$, $p = .774$. These results indicate that when accounting for the shared variance between direct and indirect item association, we no longer observe an effect of direct item association.

We also conducted an analogous linear regression to examine the relationship between w and participants' behRSA matrix components. The model included 161 observations and had an R^2 of .254 and an adjusted R^2 of .240. The results indicated a significant positive relationship between w and the indirect item association component of participants' behRSA, $b = 0.100$, $SE = 0.024$, $t(157) = 4.224$, $p < .001$. There were no significant relationships between w and direct item association, $b = 0.037$, $SE = 0.025$, $t(157) = 1.492$, $p = .138$. In this model, visual co-occurrence had a significantly negative relationship with w, $b = -0.048$, $SE = 0.018$, $t(157) = -2.446$, $p = .016$. This contrasts with the correlational version of the analysis that only found a trending negative relationship between visual co-occurrence and w. We take this to potentially stem from the fact that the teleporters that share the same first-stage state lead to completely independent reward outcomes so treating them as very dissimilar may be helpful. Knowing that RL model parameters are often correlated, we ran further linear models that are reported in the Linear Models in the online supplemental materials. Their outcomes do not qualitatively change the results reported here. Ultimately, these analyses further bolster the relationship between performance, model-based control, and abstracted cognitive maps that can be used to plan toward goals.

In sum, the encoding of higher-order representations of task structure—such as a first-stage item to its second-stage counterpart, or two

[1] Pearson's r was preregistered for these analyses, but results do not differ when using Spearman's ρ [nonparametric].

first-stage items that lead to the same second-stage item—enabled goal-directed planning toward goals, whereas representations of low-level features such as item co-occurrence did not. In our regression analyses, only the indirect item association is a positive predictor of task performance and planning. This seems to be due to collinearity in the direct item and indirect item association portions of participants' cognitive maps. Furthermore, in the case of our regression analyses representation of low-level features such as item co-occurrence seems to hinder planning. Taken together, these results suggest that the behRSA approach captures the construction of internal representation of the transition structure, the cognitive map, and the individual differences in its fidelity.

Motivation Affects Representational Change

In order to test the effect of motivation on cognitive map construction, we tested whether people used differential amounts of model-based control during the different stake conditions. Replicating prior work (Bolenz et al., 2019; Kool et al., 2017; Patzelt et al., 2019; Smid et al., 2023) model-based control was increased on high-stakes trials (mean $w_{\text{high}} = 0.56$) compared to low-stakes trials ($w_{\text{low}} = 0.54$), $F(1, 160) = 6.83$, $p = .008$, see Methods in the online supplemental materials: Stake × Arm ANOVA.[2]

Importantly, two first-stage states were associated with a high probability of a high-stakes trial (80%; high-stakes context), and the other with a low probability (20%; low-stakes context). We sought to understand whether the difference in incentives between contexts affect the representation of task structure. We predicted that aspects of the task related to higher-order structure—both direct item associations and indirect item associations—would be more strongly encoded for the high-stake compared to the low-stake context items (Figure 1A).

To test this, we ran new multiple linear regressions, estimating the coefficients separately for each stake context. We found no difference between the high- and low-stake context representations of visual co-occurrence, $t(160) = 0.45$, $p = .65$, $d = 0.04$ (Figure 3B). We found a nonsignificant trend toward a context-driven difference in the direct item-outcome representations, $t(160) = 1.74$, $p = .08$, $d = 0.14$. However, the indirect item associations were encoded more strongly for the high-stakes context compared to the low-stake context, $t(160) = 3.22$, $p = .0015$, $d = 0.25$. These results indicate that incentives lead to stronger representations for goal-directed information in task structure.

Memory for Object-Background Pairings Is Better in Higher-Stakes Contexts

Finally, we predicted that enhanced encoding of the task structure in the higher-stakes context would also enhance encoding of peripheral elements of trials in this condition. Therefore, we tested participants' memory of the object-background pairings encountered in the first stage of the two-step task. Some of these pairings were indeed encountered before (targets), but others consisted of two first-stage objects on a new exemplar from the correct background category (lures), or two first-stage objects on a background from a different first-stage state (mismatch). The mismatch trials probed highly specific memory for episodically bound object-scene pairs, because participants had encountered all of that information during the task. For each trial, participants indicated whether they had seen that combination of object and background before (Figure 6).

To test this, we computed d' sensitivity scores separately for the lure and mismatch trials. We found no effect of incentive condition for the lure trials ($t = 0.82$, $p = .4152$). For the mismatch condition, however, discrimination was higher for high-stakes compared to low-stakes trials ($t = 2.29$, $p = .0234$; Figure 3). This suggests that components of the task—even for elements that were incidental to maximizing reward—were better learned in high-stakes contexts. In particular, incentives drove enhanced memory for highly specific item-in-context information.

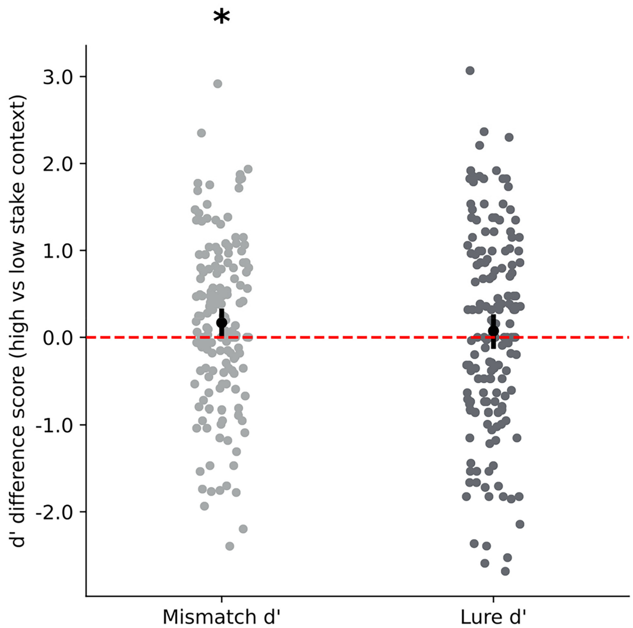
Discussion

Understanding the cognitive mechanisms underlying goal-directed planning is fundamental to the study of human behavior (Botvinick & Toussaint, 2012; Daw et al., 2011; Decker et al., 2016; Dolan & Dayan, 2013; Mattar & Lengyel, 2022; Tolman, 1948; van Opheusden et al., 2021; Wilson et al., 2014). Here, we aimed to measure one of its key components: the internal representation of task structure, also known as the cognitive map. We aimed to measure these representations at a large scale using a purely behavioral approach. To do this, we developed a novel behavioral variant of RSA to determine different sources of variability in how cognitive maps are constructed. Specifically, participants told us how related they thought the choice options they encountered in a sequential decision-making task were. Next, we correlated their ratings with three a priori hypotheses about how one might infer task structure. This allowed us to assess to which degree participants' cognitive maps reflected both more superficial and higher-order components of task structure. Strikingly, the principal ways in which participants' relatedness ratings varied were largely accounted for by our predicted models. Participants tracked aspects of task structure irrelevant for planning, such as which choice options merely co-occurred. More importantly, they also tracked components of the task that reflect its higher-order nature, such as which options are directly and indirectly linked through the transition structure.

Consistent with the idea that these relatedness ratings allow insights into cognitive maps, we found that participants whose ratings reflect the higher-order structure showed increased deployment of model-based control and better performance on the decision-making task. In contrast, planning-irrelevant components (such as item co-occurrence) did not relate to strategy selection or task performance. Moreover, the representations of indirectly related associations were more strongly present in contexts with increased incentives, further indicating the relevance of these cognitive maps for goal-directed control. The relationship between a first-stage item and the item encountered in the subsequent reward state is more complex than mere co-occurrence in the two-step task. However, one might nonetheless take this kind of representation to indicate that participants simply understand the task's transition structure. In contrast, the representation of indirect associations suggests not only an understanding of basic causality in the task, but also mapping out the space of possibilities in a more abstract

[2] Note that these estimates of model-based control are lower than the estimate obtained from the single-parameter model. This happens because only a quarter of the data contributes to each of the four model-based weighting parameter estimates, and therefore the prior (with a mode at 0.5) has a stronger influence on the estimates.

Figure 6
Posttask Surprise Memory Probe



Note. Participants show a difference in their mismatch d' sensitivity scores indicating that mismatch trials involving high-reward context backgrounds were easier to discriminate than mismatch trials involving low-reward context backgrounds. This effect is not seen in the lure d' sensitivity scores where low- and high-reward context backgrounds are equally discriminable. * represents $p < .05$, error bars are 95% CI. CI = confidence interval. See the online article for the color version of this figure.

sense. Importantly, this more abstract mapping is not merely a given for individuals who perform well at the task. That is, such a representation goes beyond immediate action-outcome relationships and suggests that some participants are linking across indirectly related experiences. Together, these findings suggest that structured representations of a task are used to guide planning toward decisions.

Probing the Cognitive Map Using behRSA

The last decade has seen an explosion of research on how humans exert model-based control, often relying on variants of the two-stage task and the dual-system RL model that we used here (Bolenz et al., 2019; Daw et al., 2011; Kool et al., 2016, 2017; Patzelt et al., 2019; Smid et al., 2023). Even though the insights from this field of research have been rich, this approach does not typically probe the cognitive map. First, in the majority of studies, it has either been assumed or ensured that the structure of a task was fully learned (e.g., Kool et al., 2017). Second, behavior in this class of tasks does not lend insight into the structure and task relevance of the cognitive map a participant has constructed. As a consequence, prior work has overlooked these representations as important sources of variance (but see Feher da Silva et al., 2023; Feher da Silva & Hare, 2020). Our behRSA approach complements and builds on this prior work. Despite participants having to learn the transition structure during task performance of the decision-making task and receiving minimal instruction during the behRSA task, we nonetheless found that their relatedness ratings revealed components of the task structure. Critically, these representations—particularly for abstractions about task structure—correlated with goal-directed control.

Representations of task structure may be learned through caching predictions about state transitions; a RL algorithm called the successor representation operates on this assumption. Though the successor representation has been supported by direct evidence in behavioral and neuroimaging work in humans (Momennejad et al., 2016; Russek et al., 2021) it does not account for other findings invoking a cost of structure learning (Collins, 2017). In addition to this, the successor representation does not predict the existence of "indirect" relations between items that share a second-stage goal. For example, agents may use probabilistic reverse inference (Solway & Botvinick, 2012; Solway et al., 2014) over their successor representations to infer indirect associations between choice options, allowing them to build up sophisticated cognitive maps from cheaply formed cached transition counts. This leads to the intriguing possibility that map construction occurs by switching between two distinct modes. In line with this idea, we found that participants more strongly represented goal-directed components of the task structure for regions of the transition structure where increased stakes were more likely. This suggests that indirect structure learning in our two-step task also imposed a cost, since participants became more willing to encode the full extent of the transition structure when it paid off more. This is not to say that the relationships were not encoded in the low-stakes contexts, more so that increased stakes further increased representational similarity among goal-relevant items. Future work may investigate what learning signals the brain uses to determine whether to engage in more extensive structure learning, and whether these two tradeoffs are truly parallel or rely on similar estimations of the value of model-based control.

Individual Differences in Cognitive Map Formation

Our behRSA results demonstrate that there exists substantial variability in the degree to which participants represent the structure of our decision-making tasks. Recent work by Feher da Silva and Hare suggests similar variability in a more conventional two-stage decision-making paradigm, which is reduced when participants are guided through a more thorough explanation of the structure of their task (Feher da Silva & Hare, 2020). In other words, constraining the space of possible cognitive maps increases the engagement of successful model-based control. By providing only instructions on the rules of the paradigm but leaving the exact transition structure unspecified, our behRSA technique allowed us to measure individual differences in the formation of cognitive maps. This flexibility may be particularly important when linking model-based decision-making to broader aspects of cognition. In particular, learning and reasoning over relational maps may be linked to the deployment of model-based control in naturalistic scenarios in which the structure is neither known nor instructed. In accordance with this idea, recent work by Rmus and colleagues found a link between the deployment of model-based control and general mapping ability (Rmus et al., 2021). In their work, performance in a separate graph learning task positively predicted the use of model-based control in the two-step task. One benefit from our, more direct, approach is that it probes the cognitive map underlying decisions within the task itself. One promising area for research would be to investigate whether general mapping ability explains the observed differences in higher-order structure learning, which then in turn explains reliance on model-based control. In other words, we predict that structure learning in the two-step task mediates the previously observed relationship between general mapping ability and the use of model-based control.

Limitations and Future Directions

Following the issue of variability, we would like to note that there was a fair amount of heteroskedasticity in whether participants generated task-related representational similarities. One of the reasons for this may be that some participants lost interest in providing ratings for pairs of objects as the task continued. Indeed, it should be noted that data collection occurred at the height of the coronavirus pandemic. It is also possible that we observed some form of retroactive interference of the behRSA task on the task structure. That is, after viewing and rating many of the possible pairings of objects, the strength of some of the previously observed relationships may have weakened. Finally, it is possible that some participants struggled with the relatively sparse nature of the behRSA instructions. We do note, however, that the results reported here are preregistered and replicate effects found in an initial sample, predicting task-related behavior after task completion. Interestingly, a follow-up exploratory sample of participants who were given a slight nudge during the behRSA phase (in terms of the task, how related do you think these objects are?), indeed showed stronger fits to our hypothesized model matrices, and stronger correlations between these fits and model-based control (Figures S3 and S4 in the online supplemental materials). It should be noted that we found stronger representations not just for the planning-related components, but also for the planning-unrelated co-occurrence component. This suggests that participants believed that this relationship was relevant to the task (even though by design this feature was not relevant to planning). Recent work using a value-guided construal model has highlighted that participants forget aspects of tasks that are not goal relevant (Ho et al., 2022). These findings seem to stand at odds with each other. A key difference between these investigations, though, is that participants in our task were repeatedly exposed to the planning-unrelated information, whereas the same goal-irrelevant aspects were only encountered once in the study by Ho et al. (2022). One possible explanation here is that the repetition of pairs of objects in the same first-stage states let them to be encoded through some form of statistical learning (Schapiro & Turk-Browne, 2015). Future research can further disentangle these mechanisms.

Another relationship we observed in our data is an enhancement of item-context recognition memory in high-reward contexts. On the one hand, it is reasonable to assume that a rich cognitive map would incorporate all potentially useful information in guiding decisions (Behrens et al., 2018). On the other hand, we note that the backgrounds used in this study were in fact irrelevant to participants' choices. That is, model-based decisions could be made purely on the basis of objects alone. Though a full understanding of this effect would require future work aimed at studying the relationship in detail, we speculate that reward drives stronger contextual encoding as a means of incorporating any potentially informative features into the representation. Indeed, prior studies show that reward states drive contextual coding (Wolosin et al., 2013) and modulate map-like representations (Garvert et al., 2023).

We note that there has been recent debate as to the accuracy of the dual-systems model as an accurate account of behavior both in this task and in general (Collins & Cockburn, 2020; Feher da Silva et al., 2023). While it is beyond the scope of this study to weigh into this debate, we acknowledge that ascribing model-based versus model-free systems to behavior in the task may not fully capture the scope of factors driving performance. Importantly, we used this paradigm to simply investigate the degree to which people rely on planning after learning the model of a task. We believe that the forward simulation of action in this task is sufficient to assess participants' planning behaviors using RL techniques. Indeed, previous work has used similar methods to demonstrate that participants engage in less planning (driving down w) when they decide to expend less effort for reward- and complexity-related reasons (Kool et al., 2017, 2018).

Though the behRSA approach is useful for behaviorally assessing participants' models and even useful for assessing the cost of structure learning, future work is needed to assess the underlying neural correlates that are allowing for map construction. In line with prior work (Schlichting et al., 2015; Schuck et al., 2016), we believe medial prefrontal cortex, orbitofrontal cortex, and hippocampus are the primary regions involved in building and using representations for our task. Relatedly, in this article, we have mainly demonstrated that participants are constructing and using their maps, but are unable to speak to how participants are doing this "online." Neuroimaging approaches are well suited to repeatedly probe representations of choice options. Another outstanding issue is the degree to which neural and behavioral RSA provide complimentary (or unique) sources of information. This comparison would provide increased understanding of the connection between reported mental and specific circuit representations, as well as the degree to which these representations are explicitly accessible.

Summary

Our findings demonstrate that it is possible to behaviorally assess the features of participants' cognitive maps. Higher-order features of these cognitive maps predict the use of model-based control and performance in an established decision-making task. Finally, the quality of these abstract cognitive maps is enhanced by the presence of higher reward, indicating that they are flexibly built and used in pursuit of goals.

Constraints on Generality

This article, and the supplement, report results from three studies. For the study reported in the main text, participants were recruited on Amazon Mechanical Turk using CloudResearch. For the pilot study and the "instruction change" study, we recruited participants from the Washington University in St. Louis subject pool. Therefore, we demonstrated the existence of these effects in a university participant pool, as well as in a broader sample of healthy younger adults in the United States. We have no reason to believe that the results depend on the specific materials or the task context, or that they would not replicate in other sensory domains. It is possible that the cognitive mechanisms producing these effects are altered in more diverse groups of people, such as during development, across aging, or in more neurodiverse samples. This provides interesting avenues for future research.

References

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, *100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002

Bolenz, F., Kool, W., Reiter, A. M., & Eppinger, B. (2019). Metacontrol of decision-making strategies in human aging. *eLife*, *8*, Article e49154. https://doi.org/10.7554/eLife.49154

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

Collins, A. G. E. (2017). The cost of structure learning. *Journal of Cognitive Neuroscience*, *29*(10), 1646–1655. https://doi.org/10.1162/jocn_a_01128

Collins, A. G. E., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience*, *21*(10), 576–586. https://doi.org/10.1038/s41583-020-0355-6

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468. https://doi.org/10.1126/science.aaf0941

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

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological Science*, *27*(6), 848–858. https://doi.org/10.1177/0956797616639301

Dimsdale-Zucker, H. R., & Ranganath, C. (2018). Chapter 27—Representational similarity analyses: A practical guide for functional MRI applications. In D. Manahan-Vaughan (Ed.), *Handbook of in vivo neural plasticity techniques* (Vol. 28, pp. 509–525). Elsevier. https://doi.org/10.1016/B978-0-12-812028-6.00027-6

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

Drummond, N., & Niv, Y. (2020). Model-based decision making and model-free learning. *Current Biology*, *30*(15), R860–R865. https://doi.org/10.1016/j.cub.2020.06.051

Feher da Silva, C., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*, *4*(10), 1053–1066. https://doi.org/10.1038/s41562-020-0905-y

Feher da Silva, C., Lombardi, G., Edelson, M., & Hare, T. A. (2023). Rethinking model-based and model-free influences on mental effort and striatal prediction errors. *Nature Human Behaviour*, *7*, 956–969. https://doi.org/10.1038/s41562-023-01573-1

Garvert, M. M., Saanum, T., Schulz, E., Schuck, N. W., & Doeller, C. F. (2023). Hippocampal spatio-predictive cognitive maps adaptively guide reward generalization. *Nature Neuroscience*, *26*(4), 615–626. https://doi.org/10.1038/s41593-023-01283-x

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806. https://doi.org/10.1038/nature03721

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. https://doi.org/10.1038/s41586-022-04743-9

Hsu, N. S., Schlichting, M. L., & Thompson-Schill, S. L. (2014). Feature diagnosticity affects representations of novel and familiar objects. *Journal of Cognitive Neuroscience*, *26*(12), 2735–2749. https://doi.org/10.1162/jocn_a_00661

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, *27*(45), 12176–12189. https://doi.org/10.1523/JNEUROSCI.3761-07.2007

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When does model-based control pay off? *PLoS Computational Biology*, *12*(8), Article e1005090. https://doi.org/10.1371/journal.pcbi.1005090

Kool, W., Gershman, S., & Cushman, F. (2018). Planning complexity registers as a cost in metacontrol. *Journal of Cognitive Neuroscience*, *30*(10), 1391–1404. https://doi.org/10.1162/jocn_a_01263

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science*, *28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Neuroscience*, *2*(1), Article 4. https://doi.org/10.3389/neuro.01.016.2008

Mattar, M. G., & Lengyel, M. (2022). Planning in the brain. *Neuron*, *110*(6), 914–934. https://doi.org/10.1016/j.neuron.2021.12.018

Momennejad, I., Russek, E., Cheong, J., Botvinick, M., Daw, N., & Gershman, S. (2016). *The successor representation in human reinforcement learning* [Preprint]. Neuroscience. https://doi.org/10.1101/083824

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press. https://books.google.com/books?id=trR-AAAAIAAJ

Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature Neuroscience*, *24*(9), 1292–1301. https://doi.org/10.1038/s41593-021-00916-3

Patil, I., Zucchelli, M. M., Kool, W., Campbell, S., Fornasier, F., Calò, M., Silani, G., Cikara, M., & Cushman, F. (2021). Reasoning supports utilitarian resolutions to moral dilemmas across diverse measures. *Journal of Personality and Social Psychology*, *120*(2), 443–460. https://doi.org/10.1037/pspp0000281

Patzelt, E. H., Kool, W., Millner, A. J., & Gershman, S. J. (2019). Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biological Psychiatry*, *85*(5), 425–433. https://doi.org/10.1016/j.biopsych.2018.06.018

Rmus, M., McDougle, S. D., & Collins, A. G. (2021). The role of executive function in shaping reinforcement learning. *Current Opinion in Behavioral Sciences*, *38*, 66–73. https://doi.org/10.1016/j.cobeha.2020.10.003

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2021). Neural evidence for the successor representation in choice evaluation. *bioRxiv*. https://doi.org/10.1101/2021.08.29.458114

Schapiro, A., & Turk-Browne, N. (2015). Statistical learning. In A. W. Toga (Ed.), *Brain mapping* (pp. 501–506). Academic Press. https://doi.org/10.1016/B978-0-12-397025-1.00276-1

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), Article 8151. https://doi.org/10.1038/ncomms9151

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, *91*(6), 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447), Article eaaw5181. https://doi.org/10.1126/science.aaw5181

Smid, C. R., Kool, W., Hauser, T. U., & Steinbeis, N. (2023). Computational and behavioral markers of model-based decision making in childhood. *Developmental Science*, *26*(2), Article e13295. https://doi.org/10.1111/desc.13295

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

Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., & Botvinick, M. M. (2014). Optimal behavioral hierarchy. *PLoS Computational Biology*, *10*(8), Article e1003779. https://doi.org/10.1371/journal.pcbi.1003779

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). The MIT Press.

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A map for social navigation in the human brain. *Neuron*, *87*(1), 231–243. https://doi.org/10.1016/j.neuron.2015.06.011

Theves, S., Fernández, G., & Doeller, C. F. (2020). The hippocampus maps concept space, not feature space. *The Journal of Neuroscience*, *40*(38), 7318–7325. https://doi.org/10.1523/JNEUROSCI.0494-20.2020

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. https://doi.org/10.1037/h0061626

van Opheusden, B., Galbiati, G., Kuperwajs, I., Bnaya, Z., Li, Y., & Ma, W.-J. (2021). *Revealing the impact of expertise on human planning with a two-player board game*. PsyArXiv. https://doi.org/10.31234/osf.io/rhq5j

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081. https://doi.org/10.1037/a0038199

Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2013). Distributed hippocampal patterns that discriminate reward context are associated with enhanced associative binding. *Journal of Experimental Psychology: General*, *142*(4), 1264–1276. https://doi.org/10.1037/a0033609

Received January 23, 2023
Revision received August 1, 2023
Accepted August 9, 2023 ■