The MIT Press

REPORT

Identifying Transfer Learning in the Reshaping of Inductive Biases

Anna Székely[1,2] (iD), Balázs Török[3], Mariann Kiss[2], Karolina Janacsek[4,5]* (iD), Dezső Németh[6,7,8]* (iD), and Gergő Orbán[1]* (iD)

[1]Department of Computational Sciences, HUN-REN Wigner Research Centre for Physics, Konkoly-Thege Miklós út 29-33., H-1121, Budapest, Hungary

[2]Department of Cognitive Science, Faculty of Natural Sciences, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary

[3]Mozalearn Ltd, 6720, Szeged, Hungary

[4]Centre for Thinking and Learning, Institute for Lifecourse Development, School of Human Sciences, Faculty of Education, Health and Human Sciences, University of Greenwich, Greenwich, SE10 9LS United Kingdom

[5]Institute of Psychology, Faculty of Education and Psychology, Eötvös Loránd University, 1071 Budapest, Damjanich u. 41-43, Hungary

[6]Université Claude Bernard Lyon 1, CNRS, INSERM, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, 69500, Bron, France

[7]NAP Research Group, Institute of Psychology, Eötvös Loránd University & Institute of Cognitive Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, Budapest, 1071, Hungary

[8]Department of Education and Psychology, Faculty of Social Sciences, University of Atlántico Medio, 35017, Las Palmas de Gran Canaria, Spain

*Senior authors.

ABSTRACT

Transfer learning, the reuse of newly acquired knowledge under novel circumstances, is a critical hallmark of human intelligence that has frequently been pitted against the capacities of artificial learning agents. Yet, the computations relevant to transfer learning have been little investigated in humans. The benefit of efficient inductive biases (meta-level constraints that shape learning, often referred as priors in the Bayesian learning approach), has been both theoretically and experimentally established. Efficiency of inductive biases depends on their capacity to generalize earlier experiences. We argue that successful transfer learning upon task acquisition is ensured by updating inductive biases and transfer of knowledge hinges upon capturing the structure of the task in the inductive bias that can be reused in novel tasks. To explore this, we trained participants on a non-trivial visual stimulus sequence task (Alternating Serial Response Times, ASRT); during the Training phase, participants were exposed to one specific sequence for multiple days, then on the Transfer phase, the sequence changed, while the underlying structure of the task remained the same. Our results show that beyond the acquisition of the stimulus sequence, our participants were also able to update their inductive biases. Acquisition of the new sequence was considerably sped up by earlier exposure but this enhancement was specific to individuals showing signatures of abandoning initial inductive biases. Enhancement of learning was reflected in the development of a new internal model. Additionally, our findings highlight the ability of participants to construct an inventory of internal models and alternate between them based on environmental demands. Further, investigation of the behavior during transfer revealed that it is the subjective internal model of individuals that can predict the transfer across tasks. Our results demonstrate that even imperfect learning in a challenging environment helps learning in a new context by reusing the subjective and partial knowledge about environmental regularities.

INTRODUCTION

One striking aspect of human intelligence is the ability to acquire new knowledge in an extremely data and time-efficient manner, especially if compared to artificial learning agents. The stunning ability of humans and other animals to learn is attributed to the ability to acquire the structural regularities of the environment and utilize this knowledge when facing similar challenges. This phenomenon, called transfer learning (also known as learning to learn or meta-learning), is possible since primate perception doesn't specialize in processing a stream of unrelated stimuli but exploits the fact that the rules and regularities structuring the environment can effectively generalize to novel situations (Brady et al., 2009; Gershman & Niv, 2010). In other words, learning is an inductive process being constrained by the rules and regularities of the world (Kemp et al., 2007; Tenenbaum et al., 2006), often referred to as structures (Kemp & Tenenbaum, 2008). The structured (and often hierarchical) organization of the environment is reflected in the acquired mental representations. Consequently, learning occurs on multiple levels too: first, on the surface, when the sensory stimuli are processed and learned within the context of a learned structure, then on an abstract (or meta-) level, when hidden structures are picked up (Behrens et al., 2018; Dekker et al., 2022; del Ojo Balaguer, 2018; Harlow, 1949; Kemp & Tenenbaum, 2008; Kóbor et al., 2020; Lake et al., 2017; Mark et al., 2020; Orbán et al., 2008; Raju et al., 2022; Samborska et al., 2022; Tse et al., 2007; Whittington et al., 2018, 2022). Our goal is to understand the computations which enable this dual-level learning.

Discovery of hidden structures is critical since it has the potential to generalize across different domains and can be critical contributors to effective learning (Acuña & Schrater, 2010; Eckstein & Collins, 2021; Schulz et al., 2020). Structured organization of the environment can be identified in specific knowledge domains, such as living species are best described by a tree structure, social structures within a group of primates can be captured by group-like structures, while the recurring characteristics of crop development can be described as cycles, to name a few examples (Kemp & Tenenbaum, 2008). Since these structures are maintained as meta-level or abstract representations, when learning on a particular domain, these structure representations serve as inductive constraints (Griffiths et al., 2010; Kemp et al., 2007; Lucas & Griffiths, 2010; Mark et al., 2020; Tian et al., 2020). Importantly, these constraints both allow learning to be much more data and time-efficient and also bias learning toward existing structure representations (Gershman & Niv, 2015; Lucas et al., 2014; Mark et al., 2020). Such influences of acquired knowledge on current interpretations are formalized in a Bayesian approach of learning as the contribution of the prior. A particularly effective form of prior for generalization over tasks or domains is referred to as hierarchical prior, through which a higher-level prior is identified over an inventory of models, which captures regularities of the structure of the earlier experiences. This high-level prior is in contrast with the low-level prior that establishes expectations for the parameters of a particular model. Critically, while priors are often considered to be given for learning, in transfer learning it is this high-level prior that is learned through experience. We will refer to these high-level structure representations as inductive biases instead of priors to emphasize their role in shaping learning. Existing inductive biases dramatically speed up learning when abstract structures are congruent with a new situation (often referred to as tasks in controlled laboratory settings) to be learned (Behrens et al., 2018; Collins & Frank, 2013; Hennies et al., 2016; King et al., 2019; Kumaran et al., 2016; Lake et al., 2017; Mark et al., 2020; Tse et al., 2007; van Kesteren et al., 2010, 2012, 2013; Wang, 2021; Wu et al., 2022), on the contrary, they can also hinder learning in case of incongruent situations (Collins & Frank, 2013; Farzanfar et al., 2023; King et al., 2019; Lucas et al., 2014; Szegedi-Hallgató et al., 2017; Török et al., 2022; Wolpert & Flanagan, 2016; Wu et al.,

2022). Operating on a task with incongruent inductive biases implies the need for updating inefficient inductive biases (Török et al., 2022). These updates are essential components of transfer learning since varying the task might imply the need for updating inductive biases. Despite the essential role of these inductive bias updates within transfer learning, we have limited knowledge on how these inductive biases adapt to novel environmental regularities.

The acquisition of hidden task structures has been most widely investigated in animal research, a topic often referred to as schema research (Harlow, 1949; King et al., 2019; Tse et al., 2007; van Kesteren et al., 2010, 2012, 2013), but humans' and artificial learning agents' capability of acquiring environmental regularities has also been investigated (Collins & Frank, 2013; del Ojo Balaguer, 2018; Farzanfar et al., 2023; Gilboa & Marlatte, 2017; Mark et al., 2020; Nelli et al., 2023; Schulz et al., 2020; Whittington et al., 2018; Wu et al., 2022). The key finding of this research is that repeated experiences create a generalisable representation of the underlying structure. These generalisable representations later permit flexible remapping to a new sensory environment which enables accelerated learning on new tasks (Behrens et al., 2018; Farzanfar et al., 2023; Gilboa & Marlatte, 2017; Hennies et al., 2016; Lewis & Durrant, 2011). To acquire these generalisable structure representations, extended exposure (often several weeks) to a specific stimulus set is necessary (Török et al., 2022; Tse et al., 2007), during which general features of the task or the environment are disentangled from the sensory stimuli (Behrens et al., 2018; Farzanfar et al., 2023; Gilboa & Marlatte, 2017; Hennies et al., 2016; Lewis & Durrant, 2011). In summary, previous research which attempted to understand how generalisable representations shape learning mainly focused on the environmental characteristics acquired, maintained and transferred across tasks. However, less emphasis was given to the fact that while inductive biases do reflect environmental regularities, learning is hardly ever perfect, thus transferred inductive biases will be inherently constrained by the fallacies of the learning phase.

Based on this insight, we propose an approach to understand the computations behind transfer learning, which relies on two important characteristics: (i) the acquisition of underlying environmental structures is hardly ever perfect, (ii) learning is always a highly individual process, even within strictly controlled laboratory conditions. Characterizing transfer learning will thus be dependent on being able to characterize the idiosyncratic errors subjects made during their learning efforts. To target this gap, we opt for a tool that is able to describe incomplete subjective task representations and individual inductive biases.

In this study we used a non-trivial sequence learning paradigm to study the inductive biases associated with temporal structure of the data. Temporal inductive biases are ubiquitous in natural phenomena, language, and quite prominently, music, where these inductive biases represent the elaborate, potentially hierarchical temporal structure underlying a scene, a language, or genre. We used a sequential structure that ensured that initial inductive biases differ from the one used in the paradigm (Török et al., 2022), which made it possible to identify the contribution of original inductive biases and permitted the study of newly emerging ones. In order to make the update of inductive biases possible, we trained participants in a specific sequence for eight days (referred to as Training phase) and investigated transfer capabilities in two additional days (Trasfer phase, Alternation test Phase), in which a new sequence is introduced that still retains the high-level structure characteristic of the sequence used during the initial training. With this training schedule, we seek to distinguish two key alternatives of learning. First, when learning occurs only on the level of stimulus sequence, i.e. by acquiring increasingly more complex sequential model as training progresses (Éltető et al., 2022; Török et al., 2022). Second, when learning also occurs at a higher level, i.e. when structure-related knowledge, the inductive bias, is updated too. To avoid the assumption of perfect learning of the task, we track day-by-day gradual development of individual learning (i.e. learning curves)

and characterize subjective knowledge about the specific task by inferring individualized task representations that we term an internal model of the task. We use Cognitive Tomography (Houlsby et al., 2013; Török et al., 2022) a machine learning-based tool, to infer individualized internal models from behavioral responses. To capture sequential knowledge in the internal model, we use a nonparametric version of the hidden Markov model, the infinite hidden Markov model (iHMM) (Van Gael et al., 2008). Beam Sampling for the Infinite Hidden Markov Model). First, we identify an initial inductive bias of individuals and track the development of the internal model to assess the evolution of the contribution of the inductive bias to later stages of the experiment. Second, we demonstrate that the dominance of the initial inductive bias is highly reduced upon exposing participants to a new sequence with a structure matching initial training. Next, we show that a second internal model is established in the short time frame of the Transfer phase. We also show that the (Training phase) internal model acquired in the Training phase is not replaced, instead, an inventory of internal models is established. Finally, we demonstrate that the structure of the subjective internal model generalizes from the Training phase to the Transfer phase of the experiment, by identifying the contribution of the updated inductive biases in the internal model acquired under the Transfer phase.

METHODS

Experimental Participants

25 healthy, young adults, 22 females, 3 male, aged between 18–22 were trained for the experimental paradigm. The population size ensured sampling the diversity of learning strategies while also identifying consistent strategies in subgroups. Hemisphere dominance was also recorded (i.e. which hand or leg is more dominant during special motor tasks). Nevertheless, since motor responses that were required during the current task relied on the use of two fingers on the right and two fingers on the left hand, we did not include hemisphere dominance or handedness in our analysis.

Data Collection and Training Schedule

Participants are trained for 10 weeks, one training session took place on each week. Multi-day training design is motivated by two reasons, first, learning and consolidating a non-trivial task structure takes time, second, it allows better characterization of individual differences despite potentially fluctuating individual performances. One session consists of 25 blocks of trials (except the tenth day of training which consists of 20 blocks), one block consists of 85 trials. The first five trials in each block are random, after which the 8 element long alternating sequence of random and deterministic elements (r-d-r-d-r-d-r-d) is repeated 10 times. The deterministic components of the sequence are governed by a specific rule that determines the upcoming observation (out of the four possible) with probability = 1 at each deterministic state. For more details see (Török et al., 2022).

Participants are expected to respond in the shortest possible time by pressing the corresponding button on a QWERTY keyboard according to the appearing stimuli, using their index and middle fingers. RTs for the first five random trials, for incorrect responses, RTs longer than 5000 ms, and RTs longer than the mean +3 standard deviation of the individual were removed from the data before the analysis. Furthermore, RTs shorter than 180 ms were also removed, since according to (Carpenter & Williams, 1995) these come from an alternative distribution. Incorrect responses take 11% of trials overall, while due to the 180 ms lower threshold overall 2.2% of the trials were removed (5.3% on Day 8). Longer than 5000 ms trials take less than 0.001% of the trials, RTs longer than the mean + 3 std account for 1.1% for all the trials.

Models

The computational models used in this paper were developed and formalized in (Török et al., 2022), here we introduce them only briefly. For a more detailed description see the original publication. We introduced three models: the Cognitive Tomography model, the Ground Truth model, and the Markov model. All three models we used in this paper were probabilistic models, relying on a Bayesian modeling approach of sequential observations. The models consist of two parts: the first part captures each participants' subjective internal representation regarding the task structure, while the second part links sequential inferences to response times. In particular, the sequential probabilistic models calculate predictive probabilities of the upcoming observation based on earlier observations, and these predictive probabilities are then transformed to response times through the behavioral model, for which we adopted the Linear Ascend to Threshold with Ergodic Rate (LATER) model (Carpenter & Williams, 1995). Models were assessed through their predictive performance, i.e. the corrrelation between the measured and predicted reaction times on a held out dataset. The model is focussing on capturing the internal model maintained by individuals. Performance of CT is not interpreted in absolute terms, instead the alternative models are used as reference models that implement specific assumptions about the internal model. Some of the variance not explained by CT can be accounted for by lapses, noise in the recording device, reactive inhibition (Török et al., 2017), or errors in dynamical inference, which are beyond the scope of the current model. The three alternative models differ in both their expressive power and their flexibility.

Markov Model.   The Markov model is a first-order Markov model, applying the classical Markov condition, where the prediction for the upcoming observation (y_{t+1}) depends only on the previous observation (y_t): $p(y_{t+1}|y_t)$. Thus, the Markov model is able to capture transitions between adjacent elements of a sequence but not the alternating dynamics of our stimulus sequence. The predictive performance of this model shows the extent to which our participants rely on a first-order dependence structure when doing the task; thus, decreasing Markov performance throughout training reflects decreasing maintenance of the first-order belief.

Ground Truth Model.   Through the Ground Truth model we formalise an ideal observer model: $p(y_{t+1}|y_1,y_2,...,y_1,y_t)$. The eight-state HMM is parametrized such that it corresponds to the true generative model of the sequence of stimuli. As such, the parameters of the Ground Truth model were not inferred from RTs, as opposed to Cognitive Tomography, but were set with 8 hidden states. In order to introduce some flexibility, there are two parameters that introduce noise into the dynamics and the emission distribution. These parameters are set to best describe individual uncertainty, as inferred from Cognitive Tomography parameters (for more details see (Török et al., 2022). As a result of this parameterization, the Ground Truth model is capable of capturing the stimulus sequence perfectly but lacks the flexibility to capture individual differences between participants. The predictive performance of this model shows the extent to which participants were able to acquire the underlying stimulus sequence.

Cognitive Tomography.   This model was formalised as an Infinite Hidden Markov Model (iHMM), which is a nonparametric extension of the original HMM (Van Gael et al., 2008). In an iHMM the number of latent states (s_1, s_2, ..., s_n) is not fixed but it is learned from the data along with the continuous parameters of the HMM (the transition and emission matrices). In practice, a Bayesian estimate of the posterior over plausible latent states and parameters is obtained by collecting samples from the posterior. For this, a modified version of the beam

sampler is defined (Török et al., 2022; Van Gael et al., 2008). Predictive probabilities are estimated by a Monte Carlo integral over samples. In contrast with the Ground Truth model, this model allows a hypothetically infinite number of hidden states, which enables this model to capture incomplete task representations, or faulty beliefs regarding the task structure: $\sum_{s_t, s_{t+1}} p(y_{t+1}|s_{t+1}) \cdot p(s_{t+1}|s_t) \cdot p(s_t|y_1, y_2, ... y_t)$. Since the iHMM is an extension of the Markov and the Hidden Markov model, Cognitive Tomography is able to capture the dynamics of both the Ground Truth and the Markov models. Due to its flexibility, Cognitive Tomography can also capture highly complex sequence dynamics and therefore behaviors which can't be described by previous simpler models.

Inference.   We use a custom sampling method, described in (Török et al., 2022). Briefly, we use a mixture of steps from the Hamiltonian Monte Carlo and the Gibbs sampling methods. Markov Chains use 1600 Gibbs-sampling steps (called slice sampling; Van Gael et al., 2008) and 30 Hamiltonian Monte Carlo are embedded in every single Gibbs step (for more details see Török et al. (2022)). To calculate predictive probabilities of upcoming stimuli we used four independently and randomly initialized chains, such that the last 30 samples were used from each chain to calculate predictive probabilities.

Response Model.   The generative model of response times relies on LATER (Linear Approach to Threshold with Ergodic Rate) model, which links subjective probabilities of upcoming stimuli with response times. A residual stochasticity of response times is describe by a reciprocal normal distribution ($r_n \sim Normal(\mu, \sigma)$). The LATER model uses the predictive probabilities (p_n) established by the internal model for each observation: $RT_n = \frac{\theta_0 - log(p_n)}{r_n}$, where μ, σ, θ_0 are parameters characterizing individual's response time model that were fitted for individuals and experimental sessions.

RESULTS

In order to gain insight into the computations underlying transfer learning, first, we need to train participants to learn a task distinct from earlier experiences and identify the acquisition of new knowledge (i.e., tracing individual learning curves); second, we need to identify how the acquired knowledge can be exploited when a new but related task is introduced. Critically, learning is always a highly individual process: Learned representations are specific to and vary across individuals, even if subjects are exposed to a well-controlled environment during learning (Török et al., 2022; Zhao et al., 2023). As a consequence of strong across-individual variation of learned representations, averaging hinders the understanding of how learned knowledge is specifically harnessed in the updated setting. In order to respect this individual nature of learning, we measured the evolution of the internal representation during learning in a subject-by-subject manner. We used Cognitive Tomography (Houlsby et al., 2013; Török et al., 2022), a state-of-the-art computational tool to infer subjective internal representations from behavior. To study transfer learning, we assume that when individuals are exposed to a novel task, they rely on an inductive bias, which reflects their initial assumptions about task contingencies. Learning comprises two components: 1, By integrating evidence and the inductive bias, participants learn a model specific to the actual task; 2, Participants update their inductive biases such that in a novel situation, learning can exploit updated expectations. To identify hallmarks of transfer learning, we need to define a task i) for which we are able to reliably identify the inductive bias that participants maintain, and ii) requires participants to acquire a new internal model, which substantially differs from the previously maintained inductive bias.

To satisfy the above-described criteria, we need behavioral measurements which are informative about the internal models. When human participants are exposed to a stimulus sequence that instructs fast motor reactions in each trial, response times (RTs) are informative regarding the expectations of the upcoming stimuli (Janacsek & Nemeth, 2013; Kóbor et al., 2017; Noorani & Carpenter, 2016). We associate these expectations with the predictive probabilities of a probabilistic model. These subject-specific probabilistic models are identified with the subjective internal models participants build during exposure to the sequence.

Internal Models Inferred From Response Times

We used the Alternating Serial Response Times (ASRT) task. In the ASRT paradigm participants are trained on a visual stimulus sequence of four stimuli in which the sequence alternates between random and deterministic elements. Participants (*n* = 25, see also Methods) are exposed to elements of the alternating sequence and are instructed to respond by pressing buttons associated with the identity of the observed stimulus (Figure 1a). The sequence progresses once a button has been pressed. The deterministic elements of the sequence are organized into a recurring 4-element-long sequence, which is thus interleaved with four randomly occurring stimuli (stimuli are sampled from a uniform distribution) (Figure 1b). While there is no direct incentive to learn the regularities of the sequence, participants are expected to implicitly learn the hidden sequence from observations (Farkas et al., 2024; Howard & Howard, 1997; Szegedi-Hallgató et al., 2019).

To assess the signatures of transfer learning in the ASRT task, we need to consider how inductive biases are updated through extended exposure to the task. For this, initial biases for sequential stimuli need to be identified. Humans' ability to learn implicit sequential structure of stimuli has been demonstrated by a vast number of studies (Acuña & Schrater, 2010; Howard & Howard, 1997; Janacsek & Nemeth, 2013; Kóbor et al., 2017; Malassis et al., 2018; Newport & Aslin, 2004; Schulz et al., 2020). Learning dependencies of adjacent events is a skill exploited under several natural challenges, such as language acquisition, when learning and executing motor tasks, or when processing or producing music (Janacsek & Nemeth, 2013; Kóbor et al., 2017; Newport & Aslin, 2004). In contrast, learning non-adjacent dependencies appears to be more challenging to humans (Malassis et al., 2018; Newport & Aslin, 2004), which suggests that humans in general maintain priors for structures which are biased towards first-order dependencies. In line with these observations, previous research has shown that in an ASRT task initially the internal models of participants were dominated by a first-order Markovian structure (Török et al., 2022). In summary, the appeal of the ASRT task to study transfer learning is manifold. First, sequence prediction is a ubiquitous task. Second, the structure of the ASRT task is highly different from the natural dynamics of the environment, and a long learning process takes place upon extended exposure. Third, the initial assumptions (the inductive bias) of participants has been previously characterized as an inductive bias that is relatively homogeneous over the participants. Next, the ASRT provides a rich dataset as reaction time measurement is obtained for long sequences (80 key presses in 20 blocks in a given session). Finally, the high-level structure of the task and its potential contribution to transfer can be identified with computational methods.

Our approach aims at understanding whether a generalized structure representation is created by abstracting the task structure from the particular sequence or increased performance on the task merely reflects learning the stimulus sequence. While the latter option claims that participants can acquire only sequence-specific knowledge, the previous implies that participants are able to update their inductive biases and utilize the learned knowledge when facing
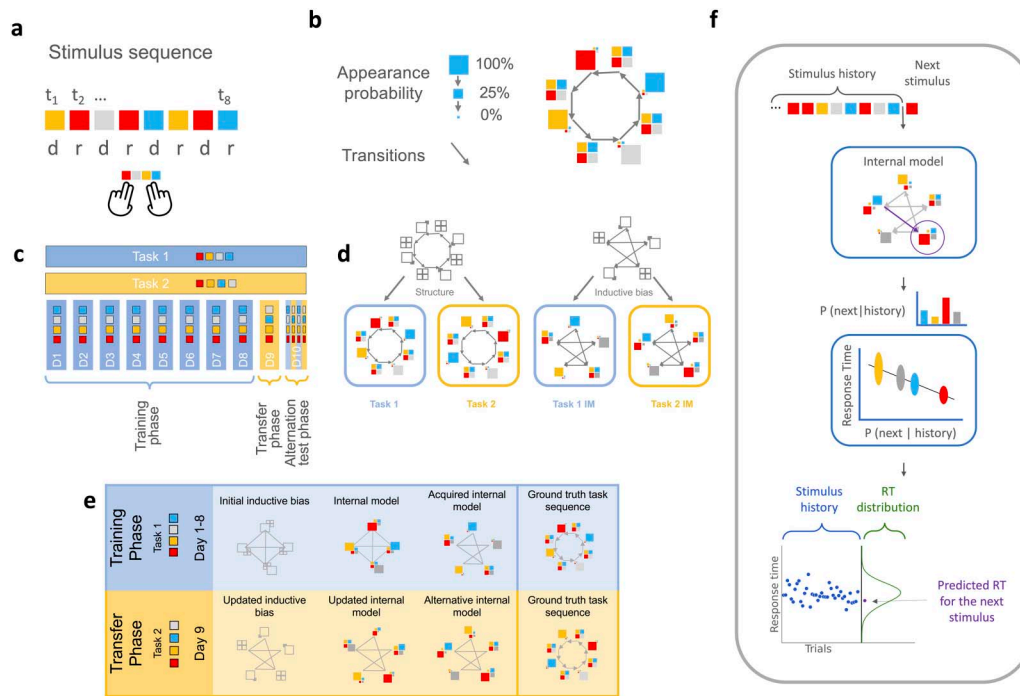
Figure 1. **Experimental paradigm and subjective internal representations. a,** An example segment of the stimulus sequence. Different colours indicate different stimuli. In a trial, one of four possible stimuli is presented, each associated with a different button that participants are required to press in order to progress the sequence. The eight-element sequence is repeated 10 times in a block. A training session consists of 25 blocks. In the eight-element sequence, random (r) and deterministic (d) elements are alternated. **b,** Representation of the probabilistic model governing the generation of the sequence. Progressing a sequence corresponds to transitions between states (black arrows). A state is characterized by the appearance probability of the four possible stimuli. The largest squares represent the deterministic states (100% appearance probability assigned to a specific stimuli, while the other 3 stimuli are associated with 0% appearance probability, represented by the smallest squares), states with four equal sized squares represent random trials (25% appearance probability associated with each stimuli). **c,** Training schedule. During the Training phase, participants are exposed for eight days (D1–D8) to a sequence governed by a given rule (blue shading, characterized by the red-yellow-gray-blue deterministic component, referred to as Task 1). During the transition phase on Day 9 of the experiment (D9) a new rule (corresponding to the red-yellow-gray-blue deterministic component, referred as Task 2) is introduced (yellow shading) that is a permutation of the rule governing the Training phase. In the Alternation phase on Day 10 of the experiment (D10) participants are exposed to both tasks, such that five-block-long sequences of the two tasks are alternated without cueing the transitions between tasks. **d,** Learning the stimulus sequence is hypothesized to occur on two levels: Learning the specific parameters that characterize each rule, and learning the underlying structure. The hierarchical nature of the task (left panel) enables us to test for these two-level computations. Under the general task structure, there are multiple possible tasks, governed by different rules. Out of the possible tasks corresponding to the structure we use two, one governing the Training phase, the other governing the Transfer phase (left panel, bottom row). Accordingly, we expect our participants to learn distinct internal models for the two rules albeit the internal models can be imperfect representations of the actual task (right panel). In the case of imperfect learning, the structure that can generalize across tasks corresponds to the number of states, transitions, and appearance probabilities characterizing the imperfect internal model. **e,** Schematics of the hypothesized acquisition of the new internal models. Top row: During the Training phase (D1–D8), learning starts with an initial inductive bias (Markovian in our case). As the training progresses, a new internal model develops, which throughout several days of training, becomes more similar to the real task structure. Bottom row: If learning entails the update of initial inductive biases, learning in the Transfer phase (D9) starts with the updated inductive bias, which supports the expedited development of a new internal model. **f,** Schematics of Cognitive Tomography. Based on stimulus history and the internal model (also inferred from RTs), appearance probabilities are calculated for the four stimuli at each trial. Given the individual RT parameters and the subjective probability of the upcoming stimuli, RTs are predicted subject-by-subject at each trial. Top box: Internal model. Colored squares indicate the probability with which any stimulus might appear in the given state, with square sizes being proportional to the appearance probabilities. Purple arrow shows momentary state transition, purple circle denotes the next state. The bar plot between the boxes shows subjective predictive probabilities for the upcoming stimulus. Middle box: Behavioral model (LATER), which relates subjective predictive probabilities to RTs. Higher subjective predictive probability will result in lower predicted RT and lower standard deviation. Bottom plot: Blue dots show RT history, the normal distribution shows the distribution of predicted RT upon the presentation of the next stimulus, and purple dot shows the RT sampled from the RT distribution.

similar but related tasks. To investigate this question, we apply a training schedule in which participants are trained for eight days (Training phase) using a stimulus sequence governed by one particular rule for the four stimuli in the deterministic component of the stimulus (Task 1), then on the ninth day (Transfer phase) we introduce a new rule to the deterministic component (Task 2) but retaining the overall task structure (Figure 1c, 1d). Finally, on the tenth day of the experiment, participants are exposed to the first and the second rule in an alternating manner, switching after every fifth block of trials (on Day 10 the session consists of 20 blocks of trials, see also Supplementary Material: Model Training and Evaluation). This alternation allows us to investigate the parallel maintenance of the two sequences: If the Training phase internal model is overwritten during the Transfer phase performance deteriorates on previously learned rule. If instead a new, separate model is learned during the Transfer phase, the two internal models are expected to be used in an alternating manner following the switches in stimulus sequences.

To capture the shifts in the internal models from the initial Markovian inductive bias towards a more complex internal model (Figure 1e), we need a framework that is able to flexibly discover internal models from a wider class of possible models. We use Cognitive Tomography to infer individualized internal models and to track individual learning curves. We use solely the sequence of response times of individuals to infer their internal model. Cognitive Tomography (CT) has two model components: The first infers individual internal models from response times, the second predicts reaction times to each stimulus based on the predictive probabilities of the internal models (Figure 1f). The internal model is formalized as a non-parametric version of the Hidden Markov Model called the Infinite Hidden Markov model (iHMM). The iHMM learns hidden stochastic dynamics over the state of the process and relates observations to these hidden states. Both the transitions between states and the observations come from probability distributions. Response time then depends on the probability of the upcoming stimulus, intuitively, a more expected stimulus is associated with a smaller response time. More formally, we use the Linear Ascend to Threshold with Ergodic Rate (LATER) (Carpenter & Williams, 1995) model (Török et al., 2022). For each participant and each day, we estimate the internal models separately. This approach lets us capture the day-by-day development of participants' internal models. The goodness of inferred internal models is evaluated through their predictive performance, that is the variance explained (Pearson's r^2) in held-out RTs (Supplementary Material: Model Training and Evaluation). Note, that r^2 is calculated on a trial-by-trial basis, instead of reporting trial-type averages. While a trigram model with 64 (4^3) trial type results in close to 0.8 r^2 values (Török et al., 2022), the trial-by-trial explained variance is lower.

To characterize the initial biases and learning that unfolds after several day-long exposure to the ASRT task, we tracked the predictive performance of the CT across the eight days of the Training phase. Extended exposure to the stimulus sequence coincided with a systematic decrease in response times and an opposite trend only appeared upon the introduction of the new sequence on Day 9 (Supplementary Figure S1b, mean RT on D1 = 367.871, std = 32.8; mean RT on D8 = 281.427; std = 18.491, mean RT on D9 = 294.577, std 15.561). CT provides a tool to flexibly characterize the internal models of individual participants, but does not assess individuals against a learning objective. As a consequence, to assess learning performance we compare the CT-inferred internal model to a baseline. Along with the CT model, we also assessed the predictive performance of the Markov model, assumed to characterize the initial inductive bias, and the Ground Truth model, reflecting the actual structure of the stimulus sequence (Figure 2a). The Markov model substantially differs from the Hidden Markov Model assumed by the CT. The Markov model completely lacks internal states, which prevents it from having memory beyond the immediately preceding experience. Without internal states, the Markov model can only account for highly impoverished dynamics in
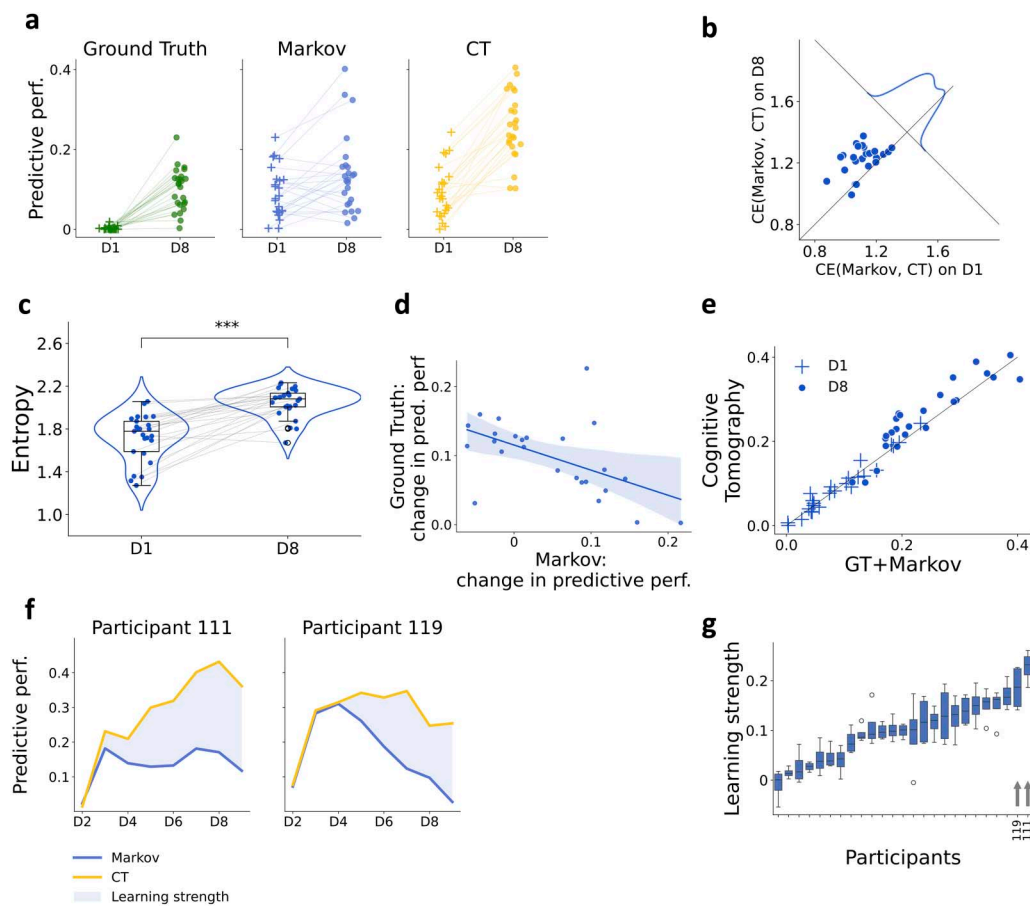
Figure 2. Characterization of learning in the Training phase. a, Predictive performance of the Ground Truth, Markov, and that of the Cognitive Tomography on the first and eighth days of training (D1 and D8, respectively). Markers indicate the predictive performance of individual participants. Initial overlap between the CT's and Markov's performance indicates that on the first day of training performance is dominated by a structure that is also captured by the Markov model. On D8 the behavior is better captured by the Cognitive Tomography than by the Markov for the majority of the participants, while the Ground Truth model also accounts for a larger amount of variance in the behavior, as opposed to the first day of training. **b,** Cross-entropy of Markov and CT models on D1 and D8. Dots show individual participants. Population level deviations from identity line show by a distribution plot. **c,** Latent state entropy on D1 and D8. Individual performances (dots) matched by a gray line across days. **d,** Change in the predictive performance of the Markov and Ground Truth model from D1 to D8. Dots show individual participants, blue shading shows the 95% confidence interval. **e,** Predictive performance of CT is proportional to the sum of variance explained by the Ground Truth and the Markov models. Markers indicate individual participants. On Day 1 (crosses) the explained variance is low, while by Day 8 (dots) predictive performance of CT reaches 0.4 for some participants. **f,** Evolution of the predictive performances of the CT and Markov models in the Training phase for two example participants. CT shows a consistent increase in predictive performance even with an eventual dramatic drop in the contribution of the Markov model. Colors of models as on panel a. Learning strength is the component of the internal model that is not explained by the Markov model (blue shading). **g,** Learning strength (averaged on D5–D8) across participants. Arrows indicate the example participants shown on c. Horizontal black lines represent the medians, boxes indicate interquartile ranges, and whiskers refer to 1.5 interquartile ranges below and above the first and third quartiles, respectively.

the sequence of observations: Direct transitions between stimuli. The Markov and Ground Truth models allow us to interpret the properties of the internal models inferred by Cognitive Tomography. Since the Markov model only accounts for immediate temporal dependencies, while the Ground Truth model relies on temporal dependencies in which the first-order statistics is not relevant, these two models account for independent sources of variance. The variances thus are additive, and the variance explained jointly by the Markov and Ground Truth models closely matches the predictive performance of the Cognitive Tomography for each participant (Figure 2e). This observation let us break down the variance explained by

Cognitive Tomography into meaningful components: The initial inductive bias and the learning strength. Since the predictive performance of the Cognitive Tomography and the Markov model are about the same level for each participant at the first days of training (Figure 2a), we identify the initial inductive bias with a Markov model. To quantify the level of match between the internal model identified by CT and the Markov model at the beginning and at the end of the Training phase, we calculated the cross-entropy between the Markov and CT models and found a consistent increase from D1 to D8 (Figure 2b, two tailed, paired *t*-test, *t*-value = −5.729, *p*-value < 0.001***, df = 24). Extended exposure to the ASRT task contributed to changes in the complexity of the inferred internal models of individuals (Supplementary Figure S1a). We assessed changes in internal model complexity by evaluating the latent variable structure of the inferred internal models. For this, we calculated the entropy of the latent states. The entropy measures the information associated with identifying a particular latent state of the internal model. This is a more intuitive measure than the number of latent states as the occupancies of the latents can be highly inhomogeneous, including extreme cases that a single state is always occupied with a number of unvisited states. The entropy was computed by calculating the distribution of latent state occupancies when data was synthesized from the internal model. Latent state entropy revealed a shift to more complex models by the end of Training phase (Figure 2c, two tailed, paired *t*-test, *t*-value = −7.058, *p*-value < 0.001***, df = 24). The predictive performance of the Ground Truth model captures how much the behavior reflects the actual stimulus sequence. Accordingly, the Ground Truth model cannot account for participant responses on D1 but the predictive performance increases considerably by D8. Variance across individuals in both the Markov and the Ground Truth model are considerable: While there is consistent increase across the population in the Ground Truth model, albeit with different intensities, there are both increases and decreases at the individual level in the Markov model (Figure 2a). We hypothesized that these changes are not independent: Those participants who acquire the Ground Truth model are more likely to rely on the Markov model later in the Training phase. Indeed, we found significant negative correlation between the changes in Markov and Ground Truth models from D1 to D8 (Figure 2d, *r* = −0.527, *p*-value = 0.007**). To gain insight into the internal model captured by CT, we investigated how much variance in response times CT accounts for and how much the Markov and Ground Truth models account for. We found that the joint predictive performance of Markov and Ground Truth models closely matches the predictive performance of the CT model both at the beginning and at the end of the Training phase (Figure 2e, *r* = 0.953, *p*-value < 0.001***). The increased amount of explainable variance on D8 compared to D1 is mainly accounted for by the acquired stimulus statistics (Figure 2a, 2d). Since the advantage of CT over the Markov model reflects the knowledge acquired about the task, we define the deviation of the predictive performance of the CT model from the Markov model as the learning strength of a participant (Figure 2f, see learning curves for all participants in Supplementary Figure S2, see also Supplementary Material: Measures for the Analyses). Thus, learning strength is a measure that characterizes the component of the predictive power of the CT, which goes beyond the Markov model. Learning strength reveals large heterogeneity of the pool of participants: Some showing negligible learning, while a larger population displays significant learning, albeit with varying levels of success during the Training phase (Figure 2g).

Learning During the Training Phase Predicts the Success of Learning in the Transfer Phase

Transfer learning is expected to facilitate learning a new task if its structure is congruent with earlier data. In terms of the ASRT task, transfer learning is expected to speed up learning of a new task if the ASRT task alternating structure of deterministic and random components is

retained in a new Task but the deterministic pattern of Task 1 is replaced with a new one in Task 2 (Figure 1d). To test this, in the Transfer phase the deterministic components of the sequence were permuted, and participants were exposed to this updated sequence on Day 9 of the experiment. While the learning strength on Day 1 didn't differ from zero (mean = 0.003, std = 0.014; one-tailed *t*-test, *t*-value = 0.181, *p*-value = 0.930, df = 24) there was a prominent difference between the predictive performance of CT and Markov models on Day 9 (mean = 0.053, std = 0.037; one-tailed *t*-test, *t*-value = 7.175, *p*-value < 0.001***, df = 24) and difference of learning strength across the days was highly significant (one-tailed, paired *t*-test, *t*-value = −6.686, *p*-value < 0.001***, df = 24) (Figure 3a). Higher learning strength on Day 9 and predictive power of learning strength between Training and Transfer phases are good indications that initial dominance of the Markov model does not return upon exposure to a new sequence. Successful transfer learning requires that upon the introduction of a new task, learning is more efficient as a result of updated inductive biases. Indeed, increased performance on D9 is consistent with this requirement. Furthermore, if the accelerated learning of the new rule on D9 emerges from the knowledge acquired during the previous eight days, varying performance on Task 1 should be reflected in individual performance level on Task 2 as well. Accordingly, performance on Task 1 predicts performance on Task 2 (*r* = 0.559, *p*-value = 0.004**) (Figure 3b). Consistent with this, we found that learning strength of participants predicted the magnitude of increase in response times between the end of the Training phase and the Transfer phase (Supplementary Figure S1c, *r* = 0.508, *p*-value = 0.010*). In contrast, the magnitude of response time reduction during the Training phase did not predict the increase between the end of the Training phase and the Transfer phase (Supplementary Figure S1d, *r* = 0.059, *p*-value = 0.779). Higher learning strength on the first day of exposure to the new task than on the first day of the initial exposure to the first task is a necessary condition for efficient transfer, but this analysis does not exclude the possibility of simply ignoring the new context while keeping the internal model acquired during the Training phase. Importantly, not only the learning strength on D8 predicts the learning strength on D9 but the contribution of the Markovian structure, which was identified with the initial inductive bias, too is correlated with acquiring a new model during the Training phase, as quantified by the learning strength on D8. A strong negative correlation is identified between the predictive performance of the Markov model on D8 and the learning strength on D8 (*r* = −0.500, *p*-value = 0.011*) (Figure 3c).
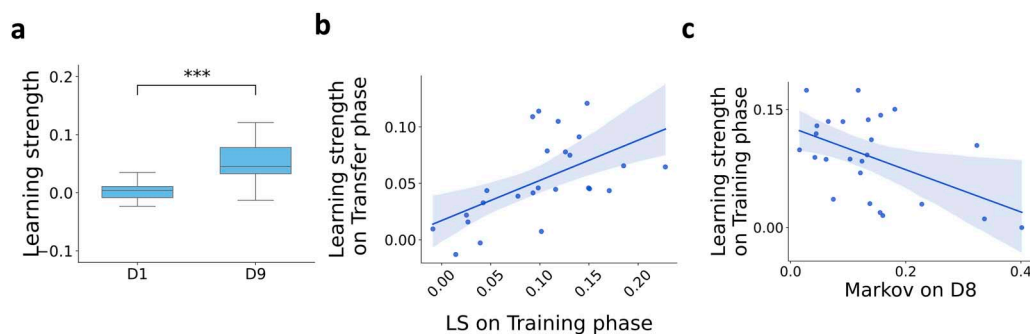

Figure 3. Characterization of learning in the Transfer phase. a, Learning strength on the first day of the Training and Transfer phases (D1 and D9, respectively). While learning strength on the first day of the Training phase was close to zero, increased performance on the first day of the Transfer phase is apparent. Horizontal black lines represent the medians, boxes indicate interquartile ranges, and whiskers refer to 1.5 interquartile ranges below and above the first and third quartiles, respectively. **b,** Learning strength on Task 1 (averaged across D5–D8) predicts learning strength on Task 2. Blue shading shows the 95% confidence interval. **c,** The Markovian structure is less dominant in the internal model of those participants who show greater learning strength. Dots show individual participants, blue shading shows 95% confidence interval.

To test whether the increased learning strength on Day 9, while exposed to Task 2, corresponds to the recruitment of a new internal model or a proliferation of the previously acquired model (Figure 4a), we trained separate models for the last day of Task 1 and for the first day of Task 2 and contrasted the predictive performance of these models. Importantly, retaining the
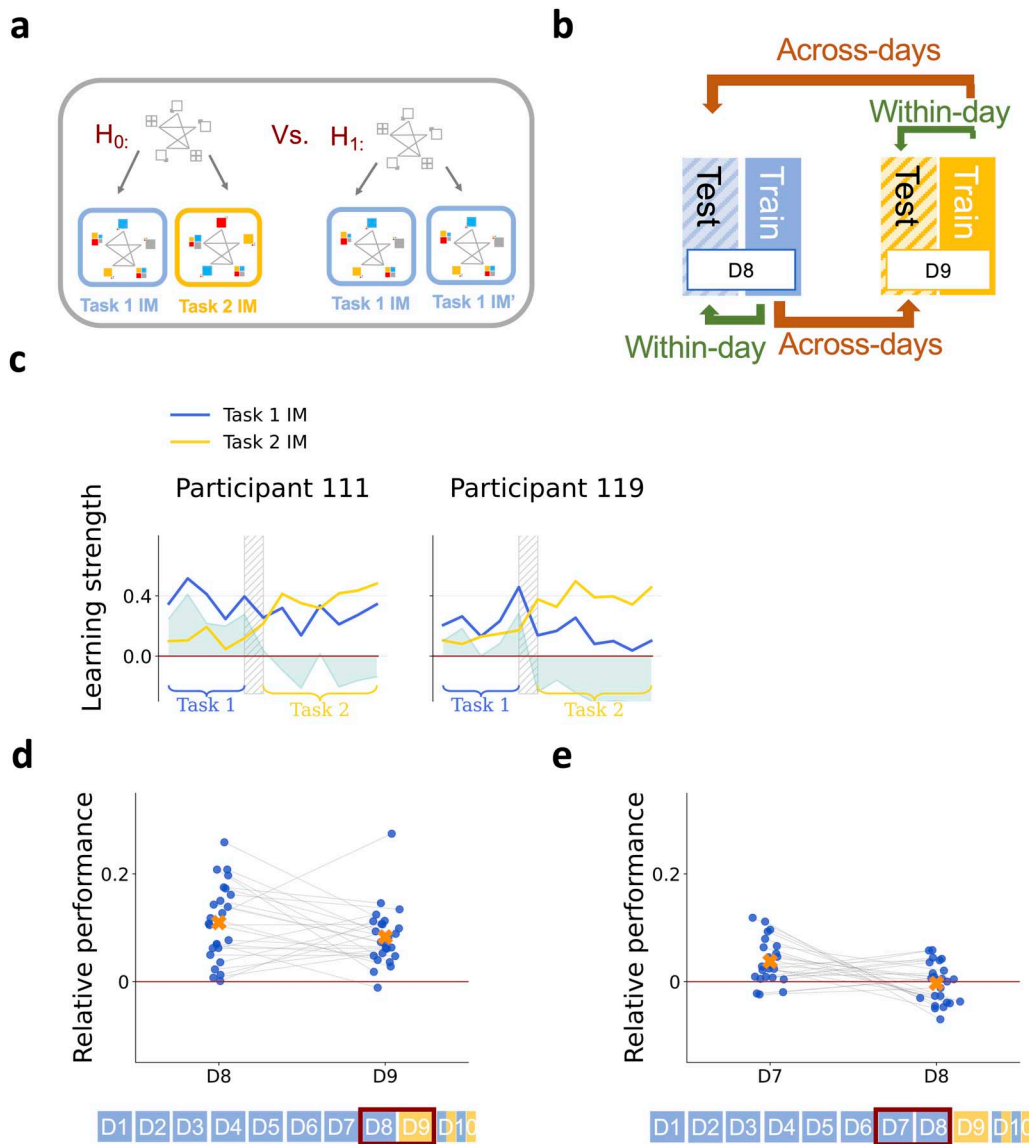

Figure 4. Acquisition of a distinct internal model for Task 2 in the Transfer phase. **a,** Alternative hypotheses for the enhanced performance in the Transfer phase (D9). H_0: Emergence of a Task 2-specific new internal model explains the enhanced predictive performance. H_1: The internal model used under Task 2 is a reflection of the proliferation of the Task 1 internal model. **b,** Schematics of inference and testing of internal models. CT is trained for each participant on D8 RTs and D9 RTs separately, then each trained model is tested on the training day (within-day, green arrows) and on the other day (across-days, brown arrows). **c,** Switching internal models for two example participants. Comparison between the within-day and across-days performance of the inferred internal models enables us to evaluate the level of dissimilarity of the two internal models. Crossover of predictive performances of D8 (blue line) and D9 (yellow line) internal models shows the switch between internal models when the task changes between D8 and D9. The extent of divergence in the two internal models reflects the dissimilar predictive power (green shading) of the two internal models on Task 1 (D8) and on Task 2 (D9). **d,** Relative performance of within-day and across-days model predictions on Day 8 and Day 9 for the pool of participants (dots). Higher values reflect stronger divergence between the internal models inferred on separate days. Orange crosses show the group mean. **e,** Same as d, but for two consecutive days during the Training phase (Day 7, and Day 8).

Task 1 internal model would mean that across-days predictions are equally strong as within-day predictions. We inferred internal models on the second half of the blocks of trials and made predictions on the early blocks of trials on the same day (within-day prediction) as well as early blocks of trials on different days (across-days prediction, Figure 4b, see also Supplementary Material: Measures for the Analyses). While proliferation of the internal model in the Transfer phase would imply strong across-days predictive performance, recruitment of a new internal model would be expected to display a crossover in the two internal models' predictive performance: Within-day predictions are expected to be strong, as opposed to a drop in predictive performance across-days (Figure 4c). We directly tested this crossover by assessing the difference between within-day and across-days predictive performances of the Task 1 and Task 2 internal models (Figure 4d). The relative performance of the within-day model was significantly above zero both on Day 8 and on Day 9 for the pool of participants (two-tailed t-test, t-value = 7.713, p-value < 0.001***; t-value = 7.500, p-value < 0.001***, df = 24, for Day 8 and Day 9 respectively). There was considerable variance across participants, in line with earlier observations about the diversity of learning strategies. To exclude the possibility that the dissimilarity of internal models reflects a non-specific deterioration of predictive performance of an internal model between days, not a change between tasks, we performed the same dissimilarity test on Day 7 and Day 8 of the Training phase. Relative performances were significantly higher for the Day 8 and Day 9 comparisons than for the Day 7 and Day 8 comparisons (anova f-value = 23.937, p-value < 0.001*** df = 24, to compare groupwise means we used the Tukey's Pairwise Group Comparisons, where all the relevant comparisons were highly significant, p-value < 0.001**). These findings show that the majority of participants were able to develop a new internal model when exposed to the new stimulus sequence, which could explain their behaviour specifically during Task 2 blocks of trials but not during Task 1 blocks.

Fast recruitment of a novel internal model during the Transfer phase indicates that learning is supported by effective inductive biases. However, the question remains open if this high level of learning enables humans to extend the initial internal model to an inventory by parallel maintaining internal models that are adequate for the Training and Transfer phases separately. If the training process entails transfer learning, the old and the new internal model coexist, otherwise, the old model is abandoned (Figure 5a). To test if parallel maintenance of internal models occurs, we extended the experiment by an additional session: On Day 10, in the Alternation phase, rules governing Task 1 and Task 2 were alternated in every five blocks of trials, without cueing the transitions (Figure 1c). If the two internal models are maintained parallel, participants are expected to alternate between them according to the stimulus governing rule. To test the parallel maintenance of the two models, we compared the predictive performance of the Task 1 and Task 2 internal models on the alternating blocks of Day 10 (Figure 5b). Participants' ability to alternate between internal models is reflected by the relative advantage of the corresponding internal model (within-task prediction) compared to the other one (across-tasks prediction; Figure 5c). The ability to alternate between the two internal models varied across participants, and this ability was, again, related to the ability to acquire new knowledge during Task 1 exposure. High learning strength on Task 1 predicted the ability to alternate between internal models according to the alternating stimuli (r = 0.473, p-value = 0.017*, Figure 5d).

Subjective Belief Transferred Across Tasks

After showing that there are distinct internal models maintained for Task 1 and Task 2, the remaining question is how we can characterize the transferred representations, i.e., the properties of the initially acquired internal model that are reused while learning in the Transfer
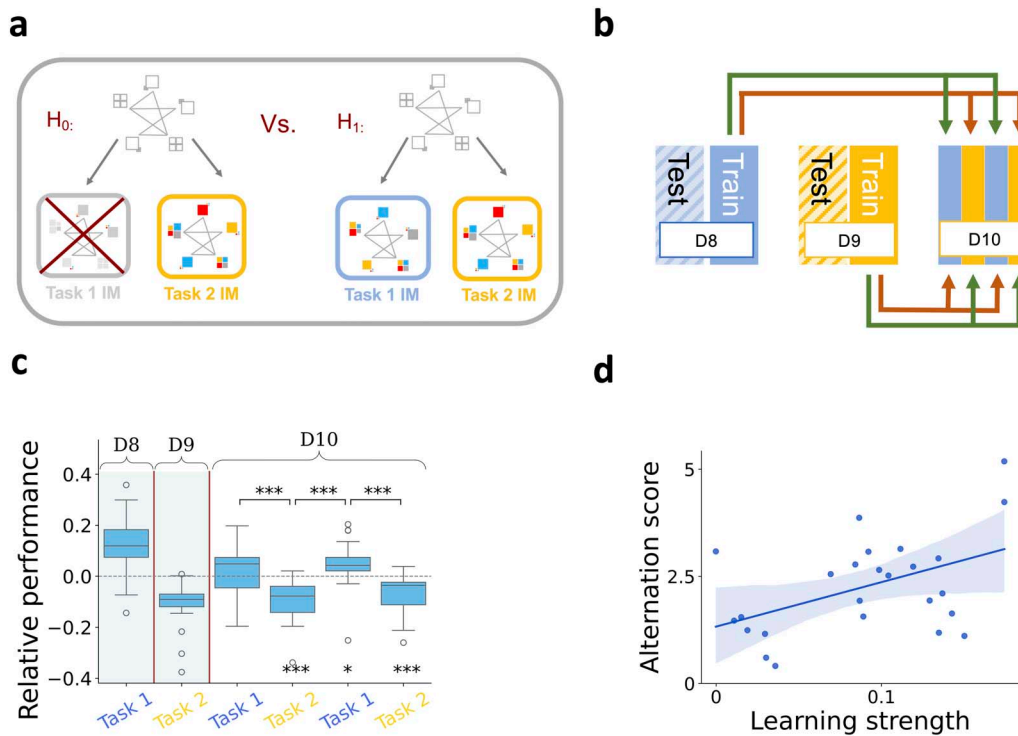
Figure 5.   Parallel maintenance of two internal models. a, Alternative hypotheses for the diverging performance of the two internal models on D8 and D9. H_0: The internal model for Task 1 was abandoned and replaced by the internal model for Task 2. H_1: Training on different stimulus governing rules entails the development of an inventory and parallel maintenance of two internal models. **b,** Schematics of inference and testing of parallel maintenance of two internal models on the alternating blocks of trials in the alternation phase (D10, stimulus governing rules alternates between Task 1 and Task 2 in five-block batches, colors as in earlier figures). CT is trained for each participant on D8 RTs and D9 RTs separately, then each trained model is tested on the corresponding (green arrows) and non-corresponding (brown arrows) task sequences of D10. **c,** Participants alternate between the two internal models according to the stimuli sequence. Relative performance is measured as a difference in the variance explained by the Task 1 internal model and Task 2 internal model. Advantage of the corresponding internal model shown by positive values on Task 1 blocks, and negative values on Task 2 blocks. The first two boxes show the within-day relative performances of D8 and D9 internal models, establishing a baseline for the expected performance measures in case of model switch. The alternating rules of the last four blocks show participants' ability to switch between internal models according to the alternating stimuli sequence on D10. **d,** Participant-by-participant (dots) assessment of the ability to alternate between two internal models as a function of learning strength during the Training phase (measured on D5–D8). Blue shading shows the 95% confidence interval.

phase. During the Training phase (Day 1–Day 8), the internal model can gradually be better characterized by the Ground Truth stimulus sequence (Figure 2a). One hypothesis of the way knowledge acquired during the Training phase can be identified in the Transfer phase is assuming that the Ground Truth model of Task 2 is tested in the Transfer phase: Since individuals show signatures of acquiring the Ground Truth structure by Day 8 (Figure 1a), the matching structure of Task 1 and Task 2 could immediately empower participants to effectively learn an internal model reflecting the properties of Task 2. (Figure 6a, H_1). However, we observed less-than-perfect learning of the ground truth sequence in individuals, relying on a combination of the initial inductive bias and the Ground Truth model even after eight days of training. As a consequence, an alternative hypothesis is that it is the internal model corresponding to the subjective belief of the participant that is reused in the new task context (Figure 6a, H_0). This internal model is referred to as subjective, since it's highly diverse across participants, and is an imperfect representation of the task structure that incorporates the characteristics of the task that a participant was able to pick up. We hypothesize that during subjective transfer, we can
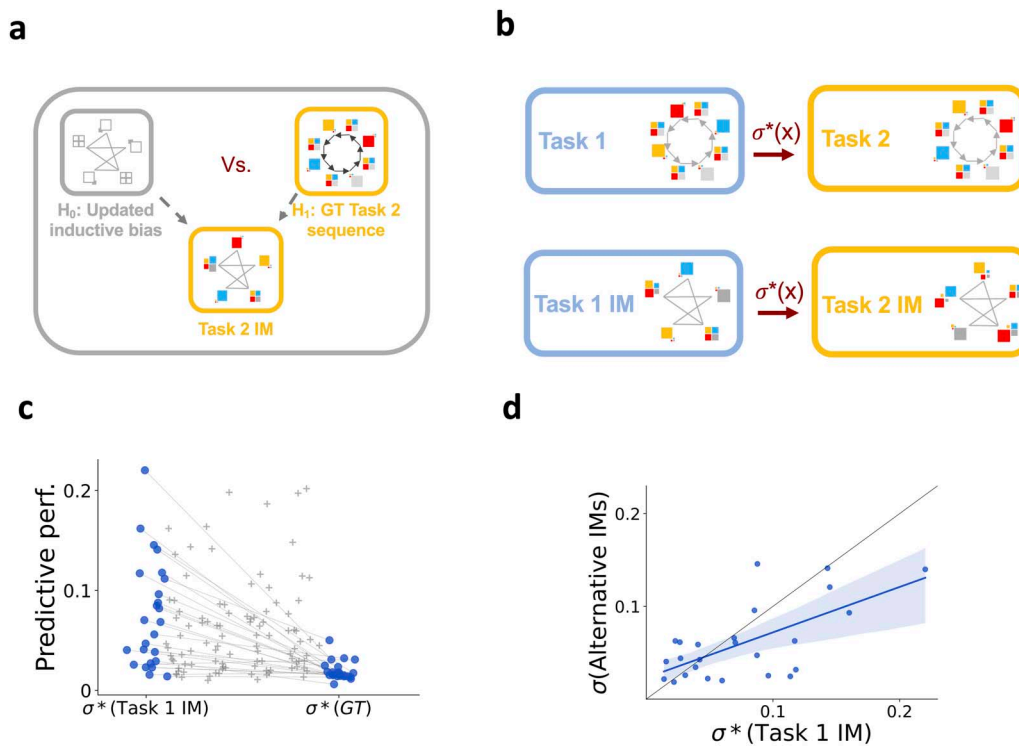
Figure 6.   Subjective belief transferred across tasks. a, Alternative hypotheses for characterizing the Task 2 internal model. H_0: The Task 2 internal model is derived through the structure of the internal model acquired during the Training phase. The structure of the Task 1 internal model determines the structure of the internal model of Task 2. H_1: The Task 2 internal model is derived from the Ground Truth sequence of Task 2. The assumption behind this hypothesis is that the gradual alignment of the internal representation of individuals with the Ground Truth Task 1 model during the Training phase enables participants to pick up the matching-structure Ground Truth model of Day 9 sequences. **b,** The Task 2 Ground Truth model can be obtained from the Task 1 Ground Truth model though a specific permutation of the possible observations associated with the states of the Task 1 model (normative permutation, σ^*). Assuming that participants have only access to their subjective belief about the 'mechanics' of the task, transfer of the acquired knowledge about the task structure entails applying the same normative permutation on the subjective internal model as the one applied to the Ground Truth model. Constructing a 'normative' internal model for Task 2 by this permutation enables us to evaluate the contribution of the subjective task representation (developed during the Training phase) to the behavior on Task 2. **c,** Comparing the predictive performance of the normative permutation of the Task 1 internal model (σ^*(Task 1 IM)) and that of the Ground Truth model. Dots show individual participants, grey crosses show alternative permutations of individual internal models. **d,** Predictive performance of the normative permutation (σ^*) of the Task 1 internal model on Task 2 response times against the predictive performance of alternative permutations of the Task 1 internal model. Black line indicates identity. Dots show the permutations for individual participants, blue line is linear regression of the data performances and shading is the 95% confidence interval.

characterize the internal model of Task 2 by applying the permutation matching the experimental manipulation to the Task 1 internal model (Figure 6b).

We contrasted these two hypotheses by calculating the predictive performance of two models: 1, The permutation of the Ground Truth model; 2, The permutation of the subjective internal model. We found that, the predictive power of the permuted subjective internal model was significantly higher compared to the predictive power of the permutation of the Ground Truth model on the Transfer phase (permuted Ground Truth model mean = 0.019, std = 0.009, permuted subjective internal model mean = 0.075, std = 0.053; two-tailed, paired *t*-test *p*-value < 0.001***, *t*-value = 5.446, df = 24) (Figure 6c). This indicates that the structure that characterizes the internal model participants acquire during the Training phase (Task 1) successfully generalizes to Task 2. Further, this transferable knowledge reflects more the characteristics of the internal model, they developed during the Training phase, than the characteristics of the Ground Truth task sequence during the Transfer phase.

We tested the specificity of this prediction by constructing alternative internal models: Instead of the 'normative' permutation motivated by the actual experimental manipulation, we constructed alternative internal models corresponding to the four alternative permutations (see Supplementary Material: Measures for the Analyses) of the original deterministic sequence. We compared the predictive performances of the normative permutation and those of alternative permutations (Figure 6d). Predictive performance of the normative permutation was significantly higher than that of the alternative permutations (t-test t-value = 5.098, p-value < 0.001***, df = 3123). Thus, our results show that the structural characteristics of the internal model acquired during the Training phase are significantly contributing to the behavior of participants in the Transfer phase. This contribution is indirect: It is a transformation of the inferred subjective internal model that predicts responses in the Training phase, a signature of reusing structural knowledge in a new context.

DISCUSSION

In this paper, our goal was to characterize the computations that facilitate transfer learning. Transfer learning requires the acquisition of a representation that is abstract enough to generalize to related but novel tasks and situations. We investigated the acquisition and exploitation of these task representations. For this, we analyzed behavioral responses in a task that is dissimilar enough from initial inductive biases so that participants needed to update their inductive biases. This enabled us to investigate specifically the contribution of that newly acquired inductive bias under a novel task. We showed that i) eight days of training is sufficient to update inductive biases and incorporate a structure that is distinct from initial priors; ii) updated inductive biases enable accelerated learning of tasks analogous to the original task; iii) training on a pair of tasks contributes to building an inventory of internal models, that is to maintain multiple models that can be recruited when circumstances necessitate their use; iv) transfer of implicit knowledge across tasks could be identified through the transformation of a subjective task representation.

Learning can be understood as a two-level computation: At a lower level, within a model family, parameters of the model need to be adjusted to reflect the properties of the environment; while on a higher-level, the right structure needs to be discovered (Fiser et al., 2010; Kemp et al., 2007; Wang, 2021). Earlier studies identified this problem as a form of Bayesian model learning and demonstrated that high-level learning can be accomplished on a rich variety of data such as object properties, causal structures or judgements (Braun et al., 2010; Gershman & Niv, 2010; Kemp & Tenenbaum, 2008, 2009; Lucas et al., 2014; Lucas & Griffiths, 2010; Zhao et al., 2023). Our work is aligned with this line of research: We assumed that initially both the structure and the parameters need to be learned by participants. Notably, we investigated a special form of such learning, learning about the sequential structure of stimuli and the structure-level knowledge was interpreted as an inductive bias for acquiring novel models of sequential stimuli. Earlier analysis has revealed that the initial inductive bias that participants rely on is heavily biased towards immediate dependencies, a simple structure that can be a parsimonious explanation in everyday circumstances and abandoning this inductive bias required extensive exposure (Török et al., 2022). In this Markovian structure transitions between observations are plagued by high uncertainty, rendering the behavior to display short time scale dependencies only (Éltető et al., 2022). Our current analysis has demonstrated that once the structure and parameters had been learned, the change in the task permitted the reuse of the learned structure. Indeed, in this new context those participants were more successful who displayed stronger signatures of learning in the Training phase and the structure of the internal model we inferred from responses was indicative of the model used in the Transfer phase.

In this paper, we argued that transfer learning capabilities of humans can be understood as the acquisition of the structure of the environment, and it is the learning of a higher-level structure that enables effective generalization and faster acquisition of new regularities. Structure learning has gained strong support in a wide array of cognitive phenomena (Heald, Wolpert, et al., 2023) such as classical conditioning in humans and animals (Courville et al., 2006; Gershman et al., 2010; Gershman & Niv, 2010) visual learning in humans (Orbán et al., 2008) inductive reasoning in humans (Kemp & Tenenbaum, 2009), sensorimotor learning in human (Braun et al., 2010; Heald et al., 2021; Tian et al., 2020). From a theoretical point of view, structure learning has been implicated as the basis of transfer learning (Hassabis et al., 2017; Wang, 2021). Our paper provides support to this idea, as we have shown that a structure learning account efficiently predicts reaction times in a transfer learning setting. However, our study does not provide a process model for transfer learning since CT only captures a snapshot of the internal representation individuals maintain.

Discovering the finer details of the principles of transfer was limited by a number of factors. First, the flexibility of CT comes with a limitation too: Stable inference of the iHMM from reaction times requires ten blocks of trials (Török et al., 2022). This can be somewhat reduced, still tracking more granular processes is limited by this property. Alternative probabilistic models with more constraints to characterize the internal model can be more effective than iHMM albeit less precise or flexible. Furthermore, using other behavioral modalities, such as gaze, might provide more information on the predictive probabilities underlying actions. Second, the population of participants displayed a high level of diversity in learning strategies in the Training phase, which differences were exacerbated in the Transfer phase. Next, during the Transfer phase of the experiment CT attempts to capture an internal model that is not stable because of potential variance in the time of abandoning the Task 1 model, as well as Task 2 model is being learned in the blocks used for inference. This volatility means that assuming a single internal model during the Transfer phase is just a coarse approximation, and predictions for the one internal model inferred by CT is more limited than in the Training phase. Nevertheless, the indication that a simple transformation of the internal model was predictive to the internal model used in the new context provides a promising indication that the two levels of learning can be identified in this setting.

With respect to the acquisition of a generalizable structure, earlier considerations highlighted that a diverse set of tasks can effectively contribute to the establishment of such a structure (Behrens et al., 2018; Harlow, 1949; Samborska et al., 2022; Tse et al., 2007; Whittington et al., 2022). In our paradigm, a single task was used to establish this generalizable structure. It could be a further line of investigation of how task diversification can promote more effective transfer. We identified successful switching between internal models when the rules of the sequence change. Tracking of the actual rule is tightly related to contextual learning (Heald, Lengyel, et al., 2023). When no cue is available for an observer to establish the actual rule, they need to learn a higher-level representation, which can be thought of as a context that establishes the actual task rules (Heald et al., 2021). The fact that a subset of participants could reliably switch between Task 1 and Task 2 internal models in the Alternation phase indicates that such internal representation of context has been established by these participants.

Transfer learning has been widely studied in the machine learning literature, which might provide insight into the underlying computations that enables the analogue mental process in primate cognition. In the machine learning literature approaches can be classified into two broad classes. A training objective-driven approach focuses on how efficient transfer can be achieved. These span methods such as engineering well-transferable weights, finding a suitably flexible model architecture (Duan et al., 2009; Galanti et al., 2022; Hong et al., 2016; Parisot et al., 2022; Yang et al., 2007; Zhuang et al., 2021), and this approach has been

successfully applied to cognitive phenomena (Mark et al., 2020; Nelli et al., 2023). An alternative approach emphasizes the importance of learned representations to improve transfer learning, often relying on hierarchies and structures that could meaningfully shape meta-level task representations (Battaglia et al., 2018; Finn et al., 2017; Grant et al., 2018; Madarasz, 2022; Madarasz & Behrens, 2019, 2020; Ortega et al., 2019; Neal & Shahbaba, 2007; Srivastava & Salakhutdinov, 2013). This line of research is well aligned with our study presented here. However, a process model that would help understanding the acquisition and manipulating knowledge at the level of structures was not addressed. While such models exist in other domains (Kemp et al., 2007; Mark et al., 2020), identifying such a process model in our sequence learning setting is a subject of future research. Recently, compelling strides have been made in how the neural circuitry is contributing to these computations (Goudar et al., 2023; Whittington et al., 2020), which might establish a framework to a biological theory of transfer.

Transfer learning is considered to be a fundamental skill of humans and has been identified in studies on non-human primates as well (Harlow, 1949). Still, the computations underlying flexible transfer are not fully understood. An important insight is that a model of the environment is a critical component of an artificial agent that can reuse its knowledge in novel situations (Walker et al., 2023). Indeed, our work showed that an internal model of the task can be reused in novel situations. Thus, studying transfer in humans might provide guidance for building better and more human-like AI.

AUTHOR CONTRIBUTIONS

A.S.: Conceptualization; Formal analysis; Methodology; Software; Visualization; Writing – original draft; Writing – review & editing. G.O.: Conceptualization; Funding acquisition; Methodology; Writing – original draft; Writing – review & editing. B.T.: Data curation; Methodology; Software. M.K.: Data Curation. K.J.: Data curation; Funding acquisition. D.N.: Data curation; Funding acquisition.

FUNDING INFORMATION

ETHICS STATEMENT

Participants were required to provide written informed consent before enrollment of the experiment. Participants received course credits for taking part in the experiment. The study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary (Approval number: 30/2012) and by the research ethics committee of Eötvös Loránd University, Budapest, Hungary. The study was conducted in accordance with the Declaration of Helsinki.

DATA AND SCRIPT AVAILABILITY

Scripts for the model training: https://github.com/mzperix/asrt-beamsampling. Experimental data and scripts for the analyses and plots: https://github.com/CSNLWigner/cogtom_transfer _learning.

REFERENCES

Acuña, D. E., & Schrater, P. (2010). Structure learning in human sequential decision-making. *PLoS Computational Biology*, *6*(12), e1001003. https://doi.org/10.1371/journal.pcbi.1001003, PubMed: 21151963

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., Gulcehre, C., Song, F., Ballard, A., Gilmer, J., Dahl, G., Vaswani, A., Allen, K., Nash, C., Langston, V., ... Pascanu, R. (2018). Relational inductive biases, deep learning, and graph networks. *arXiv*, https://doi.org/10.48550/arXiv.1806.01261

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, PubMed: 30359611

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. https://doi.org/10.1037/a0016797, PubMed: 19883132

Braun, D. A., Mehring, C., & Wolpert, D. M. (2010). Structure learning in action. *Behavioural Brain Research*, *206*(2), 157–165. https://doi.org/10.1016/j.bbr.2009.08.031, PubMed: 19720086

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, *377*(6544), 59–62. https://doi.org/10.1038/377059a0, PubMed: 7659161

Collins, A. G. E., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, *120*(1), 190–229. https://doi.org/10.1037/a0030852, PubMed: 23356780

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294–300. https://doi.org/10.1016/j.tics.2006.05.004, PubMed: 16793323

Dekker, R. B., Otto, F., & Summerfield, C. (2022). Curriculum learning for human compositional generalization. *Proceedings of the National Academy of Sciences*, *119*(41), e2205582119. https://doi.org/10.1073/pnas.2205582119, PubMed: 36191191

del Ojo Balaguer, J. (2018). *Decision-making with hierarchical representations in humans* [PhD thesis]. University of Oxford.

Duan, L., Tsang, I. W., Xu, D., & Chua, T.-S. (2009). Domain adaptation from multiple sources via auxiliary classifiers. In *Proceedings of the 26th Annual International Conference on Machine Learning*, (pp. 289–296). Association for Computing Machinery. https://doi.org/10.1145/1553374.1553411

Eckstein, M. K., & Collins, A. G. E. (2021). How the mind creates structure: Hierarchical learning of action sequences. In *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society, 43*, (pp. 618–624). Cognitive Science Society. PubMed: 34964045

Éltető, N., Nemeth, D., Janacsek, K., & Dayan, P. (2022). Tracking human skill learning with a hierarchical Bayesian sequence model. *PLoS Computational Biology*, *18*(11), e1009866. https://doi.org/10.1371/journal.pcbi.1009866, PubMed: 36449550

Farkas, B. C., Krajcsi, A., Janacsek, K., & Nemeth, D. (2024). The complexity of measuring reliability in learning tasks: An illustration using the Alternating Serial Reaction Time Task. *Behavior Research Methods*, *56*(1), 301–317. https://doi.org/10.3758/s13428-022-02038-5, PubMed: 36604378

Farzanfar, D., Spiers, H. J., Moscovitch, M., & Rosenbaum, R. S. (2023). From cognitive maps to spatial schemas. *Nature Reviews Neuroscience*, *24*(2), 63–79. https://doi.org/10.1038/s41583-022-00655-9, PubMed: 36414839

Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 1126–1135).

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: From behavior to neural representations. *Trends in Cognitive Sciences*, *14*(3), 119–130. https://doi.org/10.1016/j.tics.2010.01.003, PubMed: 20153683

Galanti, T., György, A., & Hutter, M. (2022). Generalization bounds for few-shot transfer learning with pretrained classifiers. *arXiv*. https://doi.org/10.48550/arXiv.2212.12532

Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychological Review*, *117*(1), 197–209. https://doi.org/10.1037/a0017808, PubMed: 20063968

Gershman, S. J., & Niv, Y. (2010). Learning latent structure: Carving nature at its joints. *Current Opinion in Neurobiology, 20*(2), 251–256. https://doi.org/10.1016/j.conb.2010.02.008, PubMed: 20227271

Gershman, S. J., & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science, 7*(3), 391–415. https://doi.org/10.1111/tops.12138, PubMed: 25808176

Gilboa, A., & Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013, PubMed: 28551107

Goudar, V., Peysakhovich, B., Freedman, D. J., Buffalo, E. A., & Wang, X.-J. (2023). Schema formation in a neural population subspace underlies learning-to-learn in flexible sensorimotor problem-solving. *Nature Neuroscience, 26*(5), 879–890. https://doi.org/10.1038/s41593-023-01293-9, PubMed: 37024575

Grant, E., Finn, C., Levine, S., Darrell, T., & Griffiths, T. (2018). Recasting gradient-based meta-learning as hierarchical Bayes. *arXiv*. https://doi.org/10.48550/arXiv.1801.08930

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364. https://doi.org/10.1016/j.tics.2010.05.004, PubMed: 20576465

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51–65. https://doi.org/10.1037/h0062474, PubMed: 18124807

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, *95*(2), 245–258. https://doi.org/10.1016/j.neuron.2017.06.011, PubMed: 28728020

Heald, J. B., Lengyel, M., & Wolpert, D. M. (2021). Contextual inference underlies the learning of sensorimotor repertoires. *Nature*, *600*(7889), 489–493. https://doi.org/10.1038/s41586-021-04129-3, PubMed: 34819674

Heald, J. B., Lengyel, M., & Wolpert, D. M. (2023). Contextual inference in learning and memory. *Trends in Cognitive Sciences*, *27*(1), 43–64. https://doi.org/10.1016/j.tics.2022.10.004, PubMed: 36435674

Heald, J. B., Wolpert, D. M., & Lengyel, M. (2023). The computational and neural bases of context-dependent learning. *Annual Review of Neuroscience*, *46*, 233–258. https://doi.org/10.1146/annurev-neuro-092322-100402, PubMed: 36972611

Hennies, N., Ralph, M. A. L., Kempkes, M., Cousins, J. N., & Lewis, P. A. (2016). Sleep spindle density predicts the effect of prior

knowledge on memory consolidation. *Journal of Neuroscience*, *36*(13), 3799–3810. https://doi.org/10.1523/JNEUROSCI.3162 -15.2016, PubMed: 27030764

Hong, S., Oh, J., Lee, H., & Han, B. (2016). Learning transferrable knowledge for semantic segmentation with deep convolutional neural network. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3204–3212). IEEE. https:// doi.org/10.1109/CVPR.2016.349

Houlsby, N. M. T., Huszár, F., Ghassemi, M. M., Orbán, G., Wolpert, D. M., & Lengyel, M. (2013). Cognitive tomography reveals complex, task-independent mental representations. *Current Biology*, *23*(21), 2169–2175. https://doi.org/10.1016/j .cub.2013.09.012, PubMed: 24354016

Howard, J. H., Jr., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology and Aging*, *12*(4), 634–656. https://doi.org/10.1037 /0882-7974.12.4.634, PubMed: 9416632

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: Correlated or complicated? *Cortex*, *49*(8), 2001–2006. https://doi.org/10.1016/j.cortex.2013.02.012, PubMed: 23541152

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*(3), 307–321. https://doi.org/10.1111/j.1467-7687 .2007.00585.x, PubMed: 17444972

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, *105*(31), 10687–10692. https://doi.org/10.1073/pnas.0802631105, PubMed: 18669663

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review*, *116*(1), 20–58. https://doi.org/10.1037/a0014282, PubMed: 19159147

King, B. R., Dolfen, N., Gann, M. A., Renard, Z., Swinnen, S. P., & Albouy, G. (2019). Schema and motor-memory consolidation. *Psychological Science*, *30*(7), 963–978. https://doi.org/10.1177 /0956797619847164, PubMed: 31173532

Kóbor, A., Horváth, K., Kardos, Z., Nemeth, D., & Janacsek, K. (2020). Perceiving structure in unstructured stimuli: Implicitly acquired prior knowledge impacts the processing of unpredictable transitional probabilities. *Cognition*, *205*, 104413. https:// doi.org/10.1016/j.cognition.2020.104413, PubMed: 32747072

Kóbor, A., Janacsek, K., Takács, Á., & Németh, D. (2017). Statistical learning leads to persistent memory: Evidence for one-year consolidation. *Scientific Reports*, *7*(1), 760. https://doi.org/10.1038 /s41598-017-00807-3, PubMed: 28396586

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004, PubMed: 27315762

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253. https://doi.org/10 .1017/S0140525X16001837, PubMed: 27881212

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343–351. https://doi.org/10.1016/j.tics.2011.06 .004, PubMed: 21764357

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299. https://doi.org /10.1016/j.cognition.2013.12.010, PubMed: 24566007

Lucas, C. G., & Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical bayesian models. *Cognitive Science*, *34*(1), 113–147. https://doi.org/10.1111/j.1551-6709.2009 .01058.x, PubMed: 21564208

Madarasz, T. (2022). *LPI: Learned positional invariances for transfer of task structure and zero-shot planning*. International Conference on Machine Learning 2022, Workshop on Responsible Decision Making in Dynamic Environments.

Madarasz, T. J., & Behrens, T. E. (2019). Better transfer learning with inferred successor maps. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (pp. 9029–9040). Curran Associates Inc.

Madarasz, T. J., & Behrens, T. E. (2020). *Learning transferable task schemas by representing causal invariances*. International Conference on Learning Representations (ICLR) 2020, Workshop on Causal Learning for Decision Making.

Malassis, R., Rey, A., & Fagot, J. (2018). Non-adjacent dependencies processing in human and non-human primates. *Cognitive Science*, *42*(5), 1677–1699. https://doi.org/10.1111/cogs.12617, PubMed: 29781135

Mark, S., Moran, R., Parr, T., Kennerley, S. W., & Behrens, T. E. J. (2020). Transferring structural knowledge across cognitive maps in humans and models. *Nature Communications*, *11*(1), 4783. https:// doi.org/10.1038/s41467-020-18254-6, PubMed: 32963219

Neal, R. M., & Shahbaba, B. (2007). Improving classification when a class hierarchy is available using a hierarchy-based prior. *Bayesian Analysis*, *2*(1), 221–237. https://doi.org/10.1214/07-BA209

Nelli, S., Braun, L., Dumbalska, T., Saxe, A., & Summerfield, C. (2023). Neural knowledge assembly in humans and neural networks. *Neuron*, *111*(9), 1504–1516. https://doi.org/10.1016/j .neuron.2023.02.014, PubMed: 36898375

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010 -0285(03)00128-2, PubMed: 14732409

Noorani, I., & Carpenter, R. H. S. (2016). The LATER model of reaction time and decision. *Neuroscience and Biobehavioral Reviews*, *64*, 229–251. https://doi.org/10.1016/j.neubiorev.2016 .02.018, PubMed: 26915927

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, *105*(7), 2745–2750. https:// doi.org/10.1073/pnas.0708424105, PubMed: 18268353

Ortega, P. A., Wang, J. X., Rowland, M., Genewein, T., Kurth-Nelson, Z., Pascanu, R., Heess, N., Veness, J., Pritzel, A., Sprechmann, P., Jayakumar, S. M., McGrath, T., Miller, K., Azar, M., Osband, I., Rabinowitz, N., György, A., Chiappa, S., Osindero, S., ... Legg, S. (2019). Meta-learning of sequential strategies. *arXiv*. https://doi.org/10.48550/arXiv.1905.03030

Parisot, S., Esperança, P. M., McDonagh, S., Madarasz, T. J., Yang, Y., & Li, Z. (2022). Long-tail recognition via compositional knowledge transfer. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 6929–6928). IEEE. https://doi.org/10.1109/CVPR52688.2022.00681

Raju, R. V., Guntupalli, J. S., Zhou, G., Lázaro-Gredilla, M., & George, D. (2022). Space is a latent sequence: Structured sequence learning as a unified theory of representation in the hippocampus. *arXiv*. https://doi.org/10.48550/arXiv.2212.01508

Samborska, V., Butler, J. L., Walton, M. E., Behrens, T. E. J., & Akam, T. (2022). Complementary task representations in hippocampus and prefrontal cortex for generalizing the structure of problems. *Nature Neuroscience*, *25*(10), 1314–1326. https://doi.org/10 .1038/s41593-022-01149-8, PubMed: 36171429

Schulz, E., Franklin, N. T., & Gershman, S. J. (2020). Finding struc-ture in multi-armed bandits. *Cognitive Psychology*, *119*, 101261. https://doi.org/10.1016/j.cogpsych.2019.101261, PubMed: 32059133

Srivastava, N., & Salakhutdinov, R. (2013). Discriminative transfer learning with tree-based priors. In *Proceedings of the 26th Inter-national Conference on Neural Information Processing Systems.* (Vol. 2, pp. 2094–2102). Curran Associates Inc.

Szegedi-Hallgató, E., Janacsek, K., & Nemeth, D. (2019). Different levels of statistical learning - Hidden potentials of sequence learning tasks. *PLoS ONE*, *14*(9), e0221966. https://doi.org/10.1371/journal.pone.0221966, PubMed: 31536512

Szegedi-Hallgató, E., Janacsek, K., Vékony, T., Tasi, L. A., Kerepes, L., Hompoth, E. A., Bálint, A., & Németh, D. (2017). Explicit instructions and consolidation promote rewiring of automatic behaviors in the human mind. *Scientific Reports*, *7*(1), 4365. https://doi.org/10.1038/s41598-017-04500-3, PubMed: 28663547

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318. https://doi.org/10.1016/j.tics.2006.05.009, PubMed: 16797219

Tian, L. Y., Ellis, K., Kryven, M., & Tenenbaum, J. B. (2020). Learn-ing abstract structure for drawing by efficient motor program induction. In *Proceedings of the 34th International Conference on Neural Information Processing Systems* (pp. 2686–2697). Curran Associates Inc..

Török, B., Janacsek, K., Nagy, D. G., Orbán, G., & Németh, D. (2017). Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. *Journal of Experimental Psychology: General*, *146*(4), 529–542. https://doi.org/10.1037/xge0000288, PubMed: 28383991

Török, B., Nagy, D. G., Kiss, M., Janacsek, K., Németh, D., & Orbán, G. (2022). Tracking the contribution of inductive bias to indivi-dualised internal models. *PLoS Computational Biology*, *18*(6), e1010182. https://doi.org/10.1371/journal.pcbi.1010182, PubMed: 35731822

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P., & Morris, R. G. M. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76–82. https://doi.org/10.1126/science.1135935, PubMed: 17412951

Van Gael, J., Saatci, Y., Teh, Y. W., & Ghahramani, Z. (2008). Beam sampling for the infinite hidden Markov model. In *Proceedings of the 25th International Conference on Machine Learning*, (pp. 1088–1095). Association for Computing Machinery. https://doi.org/10.1145/1390156.1390293

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2010). Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *Journal of Neuroscience*, *30*(47), 15888–15894. https://doi.org/10.1523/JNEUROSCI.2674-10.2010, PubMed: 21106827

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2013). Consolidation differentially modulates schema effects on memory for items and associations. *PLoS ONE*, *8*(2), e56155. https://doi.org/10.1371/journal.pone.0056155, PubMed: 23409144

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. https://doi.org/10.1016/j.tins.2012.02.001, PubMed: 22398180

Walker, J. C., Vértes, E., Li, Y., Dulac-Arnold, G., Anand, A., Weber, T., & Hamrick, J. B. (2023). Investigating the role of model-based learning in exploration and transfer. *arXiv*. https://doi.org/10.48550/arXiv.2302.04009

Wang, J. X. (2021). Meta-learning in natural and artificial intelli-gence. *Current Opinion in Behavioral Sciences*, *38*, 90–95. https://doi.org/10.1016/j.cobeha.2021.01.002

Whittington, J. C. R., McCaffary, D., Bakermans, J. J. W., & Behrens, T. E. J. (2022). How to build a cognitive map. *Nature Neurosci-ence*, *25*(10), 1257–1272. https://doi.org/10.1038/s41593-022-01153-y, PubMed: 36163284

Whittington, J. C. R., Muller, T. H., Mark, S., Barry, C., & Behrens, T. E. J. (2018). Generalisation of structural knowledge in the hippocampal-entorhinal system. In *Proceedings of the 32nd International Conference on Neural Information Processing Sys-tems*. (pp. 8493–8504). Curran Associates Inc.

Whittington, J. C. R., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. J. (2020). The Tolman-Eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, *183*(5), 1249–1263. https://doi.org/10.1016/j.cell.2020.10.024, PubMed: 33181068

Wolpert, D. M., & Flanagan, J. R. (2016). Computations underlying sensorimotor learning. *Current Opinion in Neurobiology*, *37*, 7–11. https://doi.org/10.1016/j.conb.2015.12.003, PubMed: 26719992

Wu, S., Eltető, N., Dasgupta, I., & Schulz, E. (2022). Learning struc-ture from the ground up—Hierarchical representation learning by chunking. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*. (pp. 36706–36721). Curran Associates Inc.

Yang, J., Yan, R., & Hauptmann, A. G. (2007). Cross-domain video concept detection using adaptive SVMs. In *Proceedings of the 15th ACM International Conference on Multimedia*. (pp. 188–197) Association for Computing Machinery. https://doi.org/10.1145/1291233.1291276

Zhao, B., Lucas, C. G., & Bramley, N. R. (2023). A model of con-ceptual bootstrapping in human cognition. *Nature Human Behaviour*, *8*(1), 125–136. https://doi.org/10.1038/s41562-023-01719-1, PubMed: 37845519

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A comprehensive survey on transfer learning. *Pro-ceedings of the IEEE*, *109*(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555