Composition of simple computational tasks captures the inductive biases of animals in network models

David L. Hocker (dh148@nyu.edu)
Center for Neural Science, New York University

Christine M. Constantinople (cmc472@nyu.edu)
Center for Neural Science, New York University

Cristina Savin (cs5360@nyu.edu)
Center for Neural Science, New York University
Center for Data Science, New York University

Abstract

Although recurrent neural networks (RNN) are now ubiquitously used by brain scientists to model neural dynamics and behavior, they are not *a priori* guaranteed to mimic animals' behavioral strategies. One reason is a fundamental model mismatch: Unlike RNNs, animals are not cognitive blank slates at task start, but they have learned through extensive prior experience. We address this issue by pretraining RNNs on tasks that mimic animals' prior inductive biases, in particular with simple cognitive "kindergarten" tasks that can be combined to perform more complex tasks. Using a rich decision-making task with latent states previously used to train rats, we demonstrate that only RNNs that incorporate kindergarten tasks into their training reflect rat-like strategies. Mechanistically, we find that the dynamics of pretrained networks are richer than those obtained with other training strategies, and that these dynamics develop during kindergarten pretraining. Overall, our approach demonstrates a simple strategy for improving RNNs as models of cognition in animals, opens up interesting questions about how previous experience shapes computational strategies that animals adopt, and provides testable predictions for neural recordings.

Keywords: curriculum learning, decision making; deep learning; dynamics; RNN

Introduction

We addressed the mismatch in prior experience between animals and RNNs by creating a novel curriculum learning (CL) paradigm that explicitly trains models first on a set of useful, basic computational skills (Fig. 1). These computational "building blocks" (*e.g.* memory, inference, evidence integration, etc.) reflect previous experiences of an animal, which are then combined via behavioral shaping towards a target end goal. We term this initial phase of training as *kindergarten curriculum learning*, to highlight that these tasks are fundamental in nature, and capture general knowledge or aptitudes that an animal likely brings to any experiment. Our approach increases task difficulty over training, similar to CL. It also uses tasks with shared underlying structure to the target task, similar to meta-learning. It is unique in that it harnesses a compositional view of complex tasks, in which simpler computational elements can be trained first, with simpler means, and reused on related tasks.

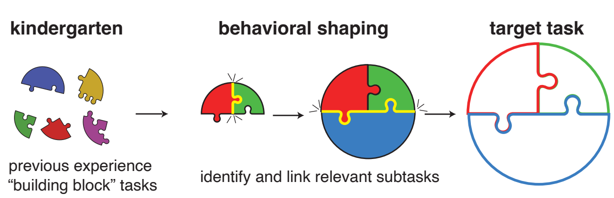

Figure 1: Kindergarten CL approach to incorporate prior experience into RNN models for complex tasks.

Methods

Task

As a complex target task, we adapted a willingness-to-wait paradigm recently studied in rats to RNNs (Mah, Schiereck, Bossio, & Constantinople, 2023) (Fig. 2A). Here, RNNs wait for reward of a known offer R, but with an unknown delay drawn from an exponential distribution (mean λ = 2.5s). RNNs can either wait for reward to arrive, or opt-out to begin a new trial. $1 - p_r = 20\%$ of offers are withheld to force the opt-out option ("catch trials"). There is a long-timescale, latent block structure to reward offers that must be inferred, which changes in an uncued fashion after an average of around 40 trials (Fig. 2B). $R = 20$ is present in each block, and behavior on those trials across blocks quantifies sensitivity to inferring the latent context. Wait times on catch trials reflect two primary features in rats and well-trained RNNs (Fig. 2C): wait times are linearly sensitive to log-reward, and are also sensitive to context, waiting longer for $R = 20$ in low blocks compared to high blocks (i.e., wait time ratio). We also modeled this task as a Markov decision process (Constantino & Daw, 2015), and found that the optimal wait time is log-linear in reward offer, and is negatively biased by the average reward earned in a block $R_{\text{av}}^{(B)}$:

$$t^* = \lambda(\log[Rp_r] - \log[R_{\text{av}}^{(B)}\lambda]). \quad (1)$$

Model and Training

We used a two-layer LSTM network to model contributions from orbitofrontal cortex (OFC) and striatum (STR), which outputs a probability of waiting π_t, and estimated state value V_t (Fig. 2D). The RNN was trained using deep meta-RL with advantage actor critic (A2C), and a policy entropy loss term to encourage exploration (Wang et al., 2018). We used deep meta-RL since it well-captures the observed, persistent representations of reward history in OFC that support trial-by-trial learning (Constantinople et al., 2019; Hocker, Brody, Savin, & Constantinople, 2021). Deep meta-RL was used to train during the wait time task; kindergarten subtasks were trained to reproduce supervised targets with their outputs o, as well as a perform a classification task with p_{block}. Inputs reflect transient trial start and offer (S_t), as well as the previous timestep's action (a_{t-1}) and reward (r_{t-1}). During final task training, kindergarten tasks acted as regularizers to our total loss function. These tasks were inspired by decomposing the wait-time task into putative sub-computations of i) working memory ii) counting elapsed time, iii) stimulus integration, and iv) state inference (Fig. 2E, top). We then built a CL sequence to first train on kindergarten tasks, followed by training stages using behavioral shaping strategies that were explicitly used in rat training, then on the target task (Fig. 2E, bottom).

Results

We found that RNNs trained with kindergarten CL captured the key aspects of rat behavior and optimal strategies: log-linear offer sensitivity, and context sensitivity (Fig. 2C, right).

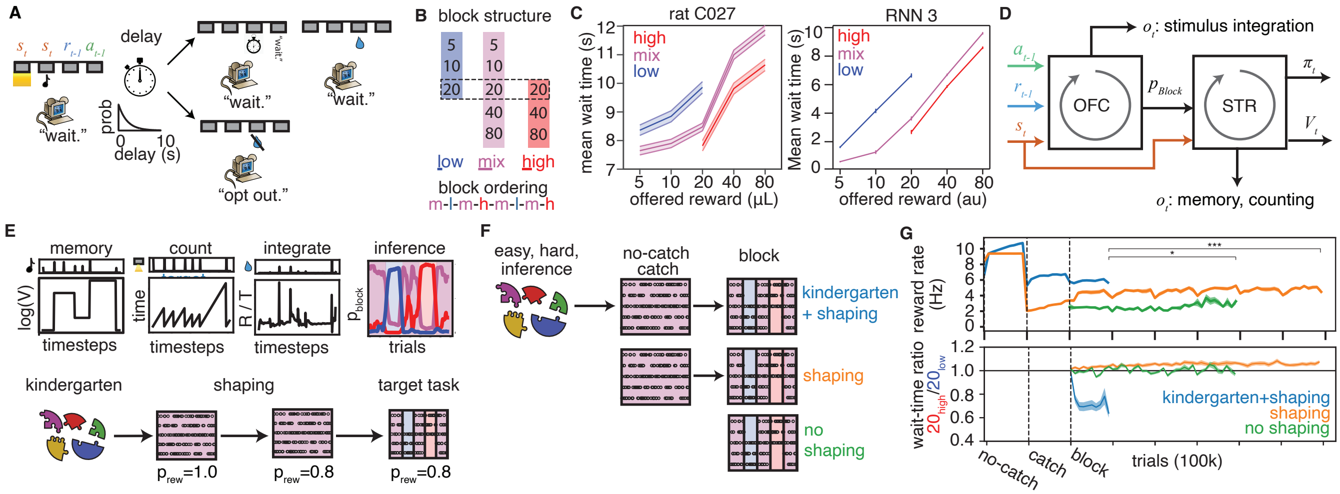

Figure 2: **A** Target wait-time task performed in RNNs. **B** Latent block structure in task. **C** Example wait time behavior on catch trials in sample rat (left) and a well-trained RNN (right). **D** Model architecture. **E** (top) Kindergarten subtasks that incorporate past experience. (bottom) Kindergarten CL training. **F** All curricula studied. **G** (top) Task performance by curriculum type, averaged across RNNs (* p$<$0.05, *** p$<$0.001, rank-sum test). (bottom) Wait time ratio over training, where < 1 reflects rat-like behavior.

Compared to other CL types (Fig. 2F), kindergarten CL achieves higher reward rates in the task, and displays the rat-like context sensitivity(Fig. **G**). Other forms of CL showed either no sensitivity to blocks, or opposite effects in which RNNs waited longer for rewards in high compared to low blocks. We also found that pretraining on irrelevant sub-tasks did not result in rat-like behavior (not shown). Overall, we found that only kindergarten CL reproduces rat behavior, suggesting that rats' inferential strategies can be captured with structured pretraining on subtasks *relevant* to target behavior.

We then studied the dynamical systems underlying well-trained RNNs. Using a PCA projection on the top two PCs (which often captured $> 90\%$ variances in many RNNs), we both empirically calculated the dynamical flow fields, as well as the linearized dynamics around fixed points and slow points. Variability existed in the geometry of the dynamics across RNNs (N=46), but overall we found a common motif in both OFC and STR dynamics. OFC dynamics trained with kindergarten CL had clear representations of blocks in which point attractors were always present in low blocks, high blocks were supported by either line or point attractors, and mixed blocks contained a saddle (Fig. 3A). Importantly, this motif was not seen in RNNs trained with simpler curricula (not shown). When conditioning STR inputs on a given block type, we found STR point attractors whose location in PC space reflected the probability of waiting (Fig. 3B). Lastly, we tracked the number of such dynamical features over training, and found that OFC dynamics had a large growth of them during their kindergarten tasks that were then consolidated in the target task (Fig. 3C), an effect recently observed in (Marschall & Savin, 2023). Finally, we found that kindergarten CL had dynamics with significantly more features than behavioral shaping alone ($p = 10^{-4}$ KS test). In summary, these results make a future testable prediction about the types of neural dynamics used to support inference-based decision making in rats.

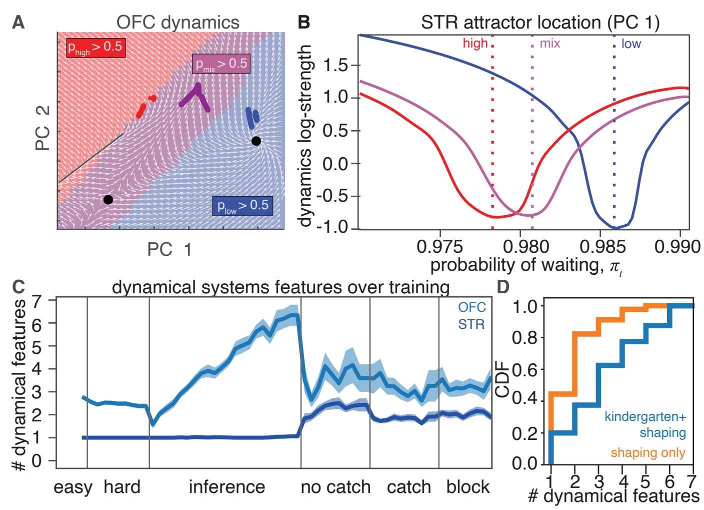

Figure 3: **A** Representative dynamics for OFC layer of well-trained RNN. **B** Similarly, for the STR layer. **C** Number of dynamical systems features over learning. **D** Comparison of dynamical systems features number in fully trained RNNs.

Conclusion

Decomposing a complex task into simpler components at the level of elements of *computation* is an essential part of behavioral shaping. RNNs can benefit from such compositional training to improve learning and to better match animal behavior. More generally, our result provide a way to model the ways in which past experiences modulate behavioral strategies.

Acknowledgments

This work was supported by NIMH grants 1K01MH132043-01A1 (DH) and 1R01MH125571-01 (CMC and CS).

References

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 837–853.

Constantinople, C. M., Piet, A. T., Bibawi, P., Akrami, A., Kopec, C., & Brody, C. D. (2019). Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *Elife*, *8*, e49744.

Hocker, D., Brody, C. D., Savin, C., & Constantinople, C. M. (2021). Subpopulations of neurons in lofc encode previous and current rewards at time of choice. *eLife*, *10*.

Mah, A., Schiereck, S. S., Bossio, V., & Constantinople, C. M. (2023). Distinct value computations support rapid sequential decisions. *Nature Communications*, *14*(1), 7573.

Marschall, O., & Savin, C. (2023). Probing learning through the lens of changes in circuit dynamics. *bioRxiv*, 2023–09.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., ... Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature neuroscience*, *21*(6), 860–868.