Compression supports low-dimensional representations of behavior across neural circuits

Dale Zhou[1] **Jason Z. Kim**[2] **Adam R. Pines**[3] **Valerie J. Sydnor**[1]
David R. Roalf[1] **John A. Detre**[1] **Ruben C. Gur**[1] **Raquel E. Gur**[1]
Theodore D. Satterthwaite[1] **Dani S. Bassett***[1,4]
[1]University of Pennsylvania [2]Cornell University [3]Stanford University [4]Santa Fe Institute
{dalezhou,pinesa,valerie.sydnor,roalf,detre,gur,raquel,sattertt}
@pennmedicine.upenn.edu, jk2557@cornell.edu, dsb@seas.upenn.edu

Abstract

Dimensionality reduction, a form of compression, can simplify representations of information to increase efficiency and reveal general patterns. Yet, this simplification also forfeits information, thereby reducing representational capacity. Hence, the brain may benefit from generating both compressed and uncompressed activity, and may do so in a heterogeneous manner across diverse neural circuits that represent low-level (sensory) or high-level (cognitive) stimuli. However, precisely how compression and representational capacity differ across the cortex remains unknown. Here we predict different levels of compression across regional circuits by using random walks on networks to model activity flow and to formulate rate-distortion functions, which are the basis of lossy compression. Using a large sample of youth ($n = 1,040$), we test predictions in two ways: by measuring the dimensionality of spontaneous activity from sensorimotor to association cortex, and by assessing the representational capacity for 24 behaviors in neural circuits and 20 cognitive variables in recurrent neural networks. Our network theory of compression predicts the dimensionality of activity ($t = 12.13, p < 0.001$) and the representational capacity of biological ($r = 0.53, p = 0.016$) and artificial ($r = 0.61, p < 0.001$) networks. The model suggests how a basic form of compression is an emergent property of activity flow between distributed circuits that communicate with the rest of the network.

1 Introduction

Brain function is metabolically expensive [1]. Neural processes can maximize efficiency by using lossy compression [2, 3], which compactly represents information with a given error [4]. Compressed representations are efficiently stored, manipulated, and communicated [5]. By discarding redundant or irrelevant information, compressed representations can foreground key features and generalizable patterns [6, 7]. Consequently, compression is a compelling model for cognitive functions that sacrifice the capacity to represent total information in order to efficiently segment and summarize relevant information. For example, working memory segments items into chunks [8, 9, 10]; emotion processing summarizes interoceptive information [11]; and value estimation summarizes promising options for action [12]. The notion that compression is functionally important for brain and artificial networks [13, 14, 15] is the *cognitive compression hypothesis*.

A key challenge to the cognitive compression hypothesis is neural evidence. One way in which compression can be implemented is by *efficient coding*, whereby the brain may reduce redundancy (e.g., activates fewer neurons) to maximize information capacity and economically use limited

Information-Theoretic Principles in Cognitive Systems Workshop at the 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

resources for behavior [2, 16]. Although many forms of compression operate by reducing redundancy, the notion of neural compression is contentious because compressive codes cannot by themselves explain observed differences in the representational capacity of lower-level visual regions versus higher-order cognitive regions [17]. Thus, a gap exists in the evidence for neural compression occurring heterogeneously across lower-level to higher-order circuits at the macroscale.

Here we address this gap by testing a network theory of compression in a large sample of youth ($n = 1,040$; ages 8 to 23 years). The theory formulates compression as a function of communication dynamics using random walks on networks. In a random walk, regions communicate with other regions by propagating signals called walkers—which represent repeated spatiotemporal motifs of activity—along structural connections to neighboring regions according to probabilities proportional to connection strengths [18]. The validity of the random walk model is supported by prior evidence that random walks on brain networks predict the strength and direction of activity [19, 20, 21], the synaptic spread of pathogens [22], the evolutionary complexity of neuroanatomy across mammalian species [23], and the speed and accuracy of human behavior [24]. We consider how random walks on networks behave according to rate-distortion functions, which are the basis of lossy compression, to predict different levels of compression in circuits [4].

To test the network theory of compression, we quantify the dimensionality and representational capacity of circuits. Low-level to higher-order circuits were sampled along a so-called *principal gradient* of organization, anchored at one end by sensorimotor functions and at the other by higher-order cognitive functions [25]. We measure the dimensionality of spontaneous (rather than evoked) circuit activity because it combines internally generated information with accurate representations of sensory stimuli [26]. Measuring dimensionality is an appropriate test of the theory because dimensionality reduction is a form of compression [27]. Due to information loss, compression also constrains the capacity to represent relevant information [28, 29]. Therefore, we measure the representational capacity of biological circuits and recurrent neural networks (RNNs) using automated meta-analysis and supervised learning for a wide range of behaviors [30, 31]. We find that circuits that support compression have low-dimensional activity and reduced representational capacity.

2 Method

Participants. We used a subset ($n = 1,040$) of an existing dataset with low in-scanner motion and high quality images [32]. Spontaneous activity was measured over 6-minute resting-state fMRI scans; motion artifact and other biological signals were removed. Structural networks were constructed from each subject's diffusion weighted imaging data [33]. Nodes of the network represented 360 cortical gray matter regions [34]. Edges of the network represented the microstructure of white matter tracts as measured by fractional anisotropy. See the **Appendix** for more details and equations.

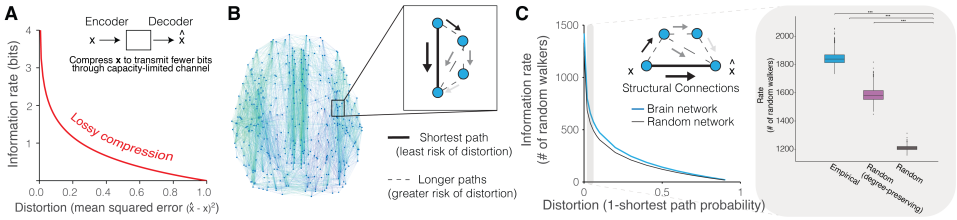

Figure 1: **Rate-distortion functions using random walks on networks.** ***(A)*** Lossy compression improves information rate efficiency at the cost of distortion. ***(B)*** In structural brain networks, we define the distortion as the probability of a random walker *not* taking the shortest path because of noise introduced by longer paths. ***(C)*** Rate-distortion functions for structural networks (blue) differ from those for random networks (black). Grey inset: at a distortion of 0.01, individual brain networks ($n = 1,040$) compress less than expected, suggesting brain networks prioritize transmission fidelity.

Rate-distortion function. We operationalized compression as the ability to send fewer random walkers across networks according to a rate-distortion function **(Figure 1)** [4]. Rate-distortion functions govern the fidelity and cost of lossy compression under an error-correcting code [35]. A random walk model naturally implements a repetition code, whereby walkers are sent repeatedly to enhance resilience to errors. We operationalize the *information rate* as the number of random walkers

sent under a repetition code. The fidelity of information transfer would then depend on whether walkers could access shortest paths because visiting more regions risks corrupting the original signal. Thus, we operationalize the *distortion* introduced by channel noise as the probability of a walker *not* taking shortest paths across the structural network. The chance to take the shortest path can be increased by sending more walkers but this enlarges the transmission. When fewer walkers can achieve the same chance to take the shortest path, we operationalized fewer walkers as compression. We count the number walkers by dividing a given probability of walkers not taking shortest paths by the probability that walkers do not take shortest paths [36, 37]. Prior research supports the validity of our operationalization, which is consistent with hypotheses about how rate-distortion functions alter with system changes, with the cost of error, and in high- or low-fidelity regimes [24, 6, 38].

Dimensionality. Dimensionality is commonly measured by the discrete number of components that are detected by eigendecomposition methods like principal components analysis and that explain a specified amount of variance. We use the *participation ratio*, which is an eigenspectrum metric that provides a continuous value of dimensionality [39, 40].

Representational capacity. We quantify the representation of 24 human behaviors across brain circuits using automated meta-analyses of neuroimaging literature [30]. This procedure provides statistical maps that associate regional activity with 24 terms that range from visual perception to emotion. Maps were validated against manual approaches and conventional maps of anatomical function. We measure representational capacity as the number of significantly associated terms. We also quantify representation of 20 multi-sensory working memory variables in RNNs trained to perform working memory tasks with supervised learning [31]. The recurrent matrix included 256 leaky units producing activity with a softplus activation function that mimics neuronal nonlinearity. We measure the representational capacity as the amount of unit activity variance across trials per task, which prior work used to quantify the amount of represented stimulus information.

3 Results

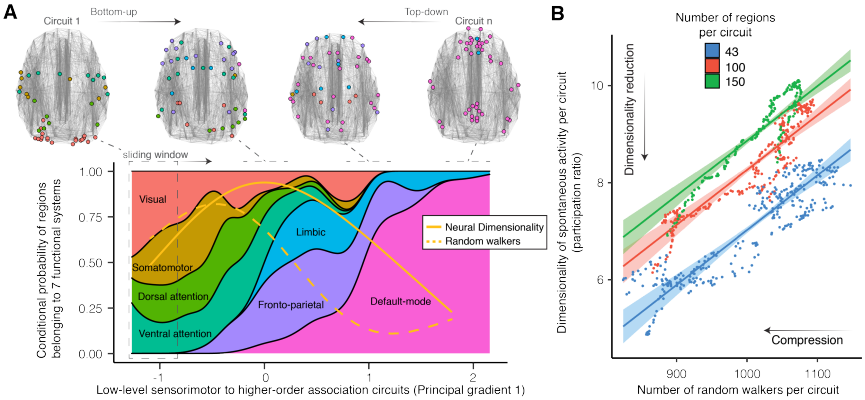

Figure 2: **Circuit compression explains the dimensionality of spontaneous activity.** ***(A)*** Dimensionality and random walkers across circuits that are sampled along the principal gradient. For visual comparison, values are normalized between 0 and 1 and reported with the conditional probability of regions belonging to 7 functional systems. ***(B)*** Circuit compression predicts the dimensionality of spontaneous activity across circuits of different sizes.

Circuit compression predicts low-dimensional spontaneous activity and reduced representational capacity. We slid a window of binned regions across the principal gradient to sample low-level to higher-order circuits. For each circuit, we measured compression and dimensionality (**Figure 2**). We tested linear and non-linear effects while minimizing over-fitting using generalized additive models, a non-parametric regression method. As predicted, circuits that compress information by using fewer random walkers had lower-dimensional activity (linear effect t=12.13, p <0.001). Compression and dimensionality are inverted-U functions along the principal gradient (non-linear effect F=350.9, p<0.001). As predicted, circuits with more compression had reduced representational capacity in humans ($r = 0.53, p = 0.016$) and in RNNs ($r = 0.61, p < 0.001$) for diverse behaviors (**Figure 3**).

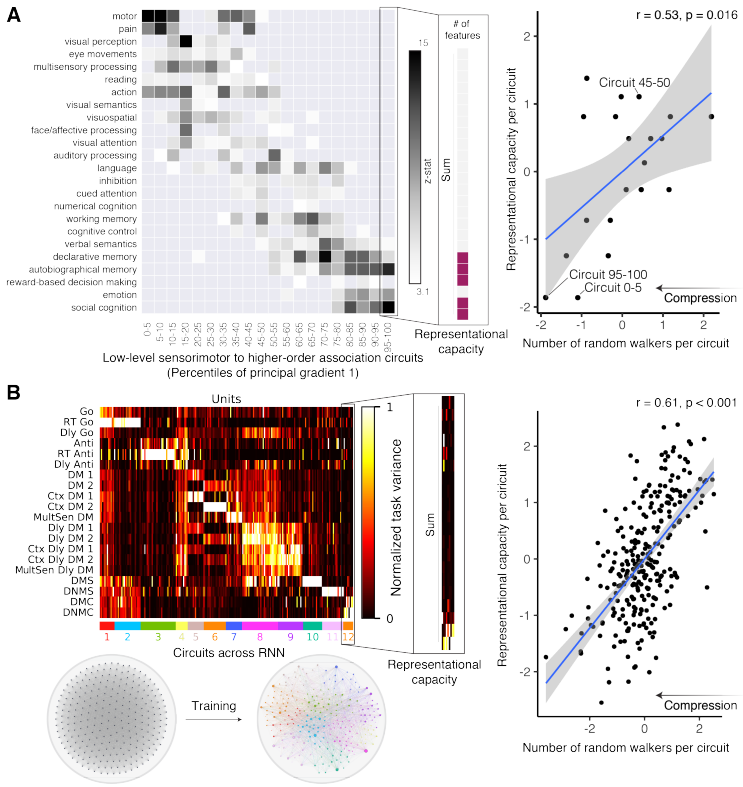

Figure 3: **Circuit compression explains the representational capacity of biological and artificial neural circuits.** ***(A) Left***: Circuits along a sensorimotor to association continuum (columns) have varying representational capacity for information relevant to diverse behaviors (rows). *Right*: Circuit compression constrains representational capacity. ***(B) Left***: Example of one RNN with 256 units (columns) that was trained with supervised learning to perform 20 multisensory working memory tasks (rows; see **Appendix** for legend). Following training, RNNs learned to use specialized circuits depicted as 12 colored clusters from t-distributed stochastic neighbor embedding. Matrix entries show different representational capacity for each task defined by the task variance. *Right*: Circuit compression constrains representational capacity. Matrices reproduced from code in [41] and [31].

4 Discussion

Consistent with the cognitive compression hypothesis, our findings link the function of higher-order cognitive circuits to compression and low-dimensional activity. We found that dimensionality and compression are related and are non-linear functions of circuit hierarchy: highest compression exists at the levels of sensorimotor and association function and lowest compression exists at the level of attention. Circuits with more compression represented fewer behaviors, an explanation of how circuits may efficiently specialize their function. Our findings relate information transmission to dimensionality, a property thought to support information processing. Expanding dimensionality improves representation and classification, whereas reducing dimensionality improves generalizability to unfamiliar contexts [42, 43, 44, 45, 29]. Heterogenous levels of dimensionality support flexible behavior and compositionality of compact representations [46, 47, 48]. To further understand information processing, the model could be used to explain how the brain generates and updates compressed models of the environment with differing constraints of dimensionality, timescales, and controllability [49, 50, 39, 45], key elements of predictive coding [51, 52, 16]. Our model admits several limitations. Although investment of metabolic resources and myelin improves a repetition code's efficiency, the code by itself is rudimentary and has limited efficiency, motivating studies of more sophisticated compression [24]. Strengths of the model include being robust at the individual level and bridging the compression of brain regions to behavior. Hence, future work could investigate if skewed levels of compression explain how circuit computations go awry in mental illnesses due to over-compression that results in inaccurate yet highly generalized models of the world.

5 Acknowledgments

D.Z. acknowledges support from the National Institute of Mental Health (F31MH126569). D.S.B. acknowledges support from the John D. and Catherine T. MacArthur Foundation, the Swartz Foundation, the Paul G. Allen Family Foundation, the Alfred P. Sloan Foundation and the NSF (IIS-1926757). D.S.B. and T.D.S. acknowledge support from the National Institute of Mental Health (R01MH113550).

A Appendix

A.1 Participants

Participants completed a cross-sectional imaging protocol. Exclusion criteria included excessive in-scanner motion, serious health conditions, and quality control models. The participants used in this paper are a subset of the 1,601 participants who completed the cross-sectional imaging protocol. We excluded participants with health-related exclusionary criteria (n = 154) and with scans that failed a rigorous quality assurance protocol for diffusion-weighted imaging (DWI; n = 162). We further excluded subjects with incomplete or poor ASL and field map scans (n = 60). Finally, participants with poor quality T1-weighted anatomical reconstructions (n = 10) were removed from the sample. The final sample contained 1042 subjects (mean age = 15.35, SD = 3.38 years; 467 males, 575 females). All adult participants provided informed consent; all minors provided assent and their parent or guardian provided informed consent.

A.2 Formulating rate-distortion functions using random walks on networks

Our formulation of the rate-distortion function uses exact expressions for the mean first passage time of random walks on networks [53]. We operationalize rate-distortion functions as the number of random walkers r_{ij} which begin at node i that were required for at least one to travel along the shortest path to another node j with probability η [36, 37]. We define the transition probability matrix as $\mathbf{U} = \mathbf{W}\mathbf{L}^{-1}$, where each entry W_{ij} describes the weight of the connection from node i to node j, and each entry of the diagonal matrix $\mathbf{L}$ is the strength of each node i given by $\sum_i W_{ij}$. Intuitively, each entry U_{ij} defines the probability of a random walker traveling from node i to node j in one step.

We operationalized distortion as the probability that one of the sent walkers does *not* take the shortest path. To obtain this probability we first need to compute the probability that a random walker travels from node i to node j along the shortest path. We define a new matrix $U'(i)$ that is equivalent to $\mathbf{U}$ but with the non-diagonal elements of row i set to zero and $U_{ii} = 1$ as an absorbant state. Then, the probability of randomly walking from i to j along the shortest path is given by $1 - \sum_{n=1}^{N} [U'(i)^H]_{in}$, where H is the number of connections composing the shortest path from i to j.

We can use these definitions to determine the number of random walkers r required to guarantee (with probability η) that at least one of them travels from i to j along the shortest path:

$$r_{ij}(\eta) = \frac{\log(1-\eta)}{\log\left(\sum_{n=1}^{N} [U'(i)^H]_{in}\right)}. \tag{1}$$

In our analyses, we calculate r_{ij} for $\eta = 0.999$ for each participant. This calculation returns a right stochastic matrix U'_i, where the number of random walkers a brain region receives is $(r_{ji}(\eta))$ averaged over j.

Finally, we substitute these definitions as our choice of the rate-distortion function. A signal x is encoded as $\hat{x}$ with a level of distortion D that depends on the information rate R. We define the distortion function of any signal x from brain region i to a compressed representation $\hat{x}$ decoded in brain region j as $d(x, \hat{x})_{ij} = (1 - \eta)$, where η denotes the probability that a walker gets from node i to node j along the shortest path. The rate-distortion function is $R(D) \equiv r_{ij}(d(x, \hat{x})_{ij})$, where the rate is $r_{ij}(\eta)$ required to achieve a tolerated level of distortion $d(x, \hat{x})_{ij}$.

A.3 Dimensionality

Principal components analysis is a common multivariate method to decompose data into a mutually orthogonal (uncorrelated) set of data ordered by the amount of variance explained [54, 55]. In principal components analysis, the principal components are the eigenvectors v, and the proportion of variance explained by the components are the eigenvalues λ of the demeaned and centered $N \times N$ covariance matrix of spontaneous activity. The participation ratio normalizes eigenvalues $\tilde{\lambda_i}$ with the equation: $\tilde{\lambda_i} = \frac{\lambda_i}{\sum_j \lambda_j}$ [39, 40]. The dimensionality then is:

$$\frac{1}{\sum_i \tilde{\lambda_i}^2}. \tag{2}$$

A.4 Imaging acquisition

Diffusion imaging data and all other MRI data were acquired on the same 3T Siemens Tim Trio whole-body scanner and 32-channel head coil at the Hospital of the University of Pennsylvania. DWI scans were obtained using a twice-focused spin-echo (TRSE) single-shot EPI sequence (TR = 8100 ms, TE = 82 ms, FOV = 240 mm^2/240 mm^2; Matrix = RL: 128/AP:128/Slices:70, in-plane resolution (x & y) 1.875 mm^2; slice thickness = 2 mm, gap = 0; FlipAngle = 90°/180°/180°, volumes = 71, GRAPPA factor = 3, bandwidth = 2170 Hz/pixel, PE direction = AP). The sequence employs a four-lobed diffusion encoding gradient scheme combined with a 90-180-180 spin-echo sequence designed to minimize eddy current artifacts. The complete sequence consisted of 64 diffusion-weighted directions with b = 1000 s/mm^2 and 7 interspersed scans where b = 0 s/mm^2. Scan time was about 11 min. The imaging volume was prescribed in axial orientation covering the entire cerebrum with the topmost slice just superior to the apex of the brain [56].

A.5 Connectome construction

Cortical gray matter was parcellated according to the Glasser atlas [34], defining 360 brain regions as nodes for each subject's structural brain network, denoted as the weighted adjacency matrix $\mathbf{A}$. DWI data was imported into DSI Studio software and the diffusion tensor was estimated at each voxel [33]. For deterministic tractography, whole-brain fiber tracking was implemented for each subject in DSI Studio using a modified fiber assessment by continuous tracking (FACT) algorithm with Euler interpolation, initiating 1,000,000 streamlines after removing all streamlines with length less than 10mm or greater than 400mm. Fiber tracking was performed with an angular threshold of 45, a step size of 0.9375mm, and a fractional anisotropy (FA) threshold determined empirically by Otzu's method, which optimizes the contrast between foreground and background [33]. FA was calculated along the path of each reconstructed streamline. For each subject, edges of the structural network were defined where at least one streamline connected a pair of nodes. Edge weights were defined by the average FA along streamlines connecting any pair of nodes. The resulting structural connectivity matrices were not thresholded and contain edges weighted between 0 and 1.

A.6 Metabolic energy

Compressed codes economically use limited metabolic resources, which is a key benefit of efficient coding [3]. Metabolic energy expenditure is tightly linked to regional cerebral blood flow [57, 58]. Therefore we normalize our predictions of compression by dividing the number of random walkers by the regional cerebral blood flow (CBF), an operationalization of metabolic energy expenditure [59, 60]. Arterial-spin labeling data was used to determine the regional cerebral blood flow (mL per 100 g tissue per second). We use this normalized metric when predicting the dimensionality of spontaneous activity. The normalized metric is in units of number of random walkers per mL per 100 g tissue per second. Sensitivity analysis suggests that our theory remains predictive of dimensionality without this normalization step (**Figure 4**). Circuits that transmit fewer random walkers to the whole brain have low-dimensional spontaneous activity.

CBF was quantified from control-label pairs using ASLtbx [61], as previously described [62]. We consider f as CBF, δM as the difference of the signal between the control and label acquisitions, R_{1a} as the longitudinal relaxation rate of blood, τ as the labeling time, ω as the post-labeling delay

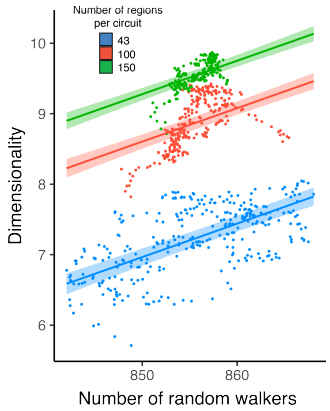

Figure 4: **Circuit compression explains the dimensionality of spontaneous activity.** ***(A)*** The number of random walkers (unnormalized for CBF) predicts the dimensionality of spontaneous activity across circuits of different sizes. Plotted with a generalized additive model, which is a non-parametric regression method.

time, α as the labeling efficiency, λ as the blood/tissue water partition coefficient, and M_0 as the approximated control image intensity. Together, CBF f can be calculated according to the equation:

$$f = \frac{\Delta M \lambda R_{1a} \exp\left(\omega R_{1a}\right)}{2 M_0 \alpha} \left[1 - \exp\left(-\tau R_{1a}\right)\right]^{-1}. \tag{3}$$

Because prior work has shown that the T1 relaxation time changes substantially in development and varies by sex, this parameter was set according to previously established methods, which enhance CBF estimation accuracy and reliability in pediatric populations [63, 64].

A.7 Principle gradient 1

The principal gradient of intrinsic functional connectivity reflects a functional processing hierarchy from sensory cortex to association cortex. We used both a publicly available atlas of the principal gradient and published tools to calculate the principal gradient in our own data [65]. Briefly, these tools use a form of Lacplacian eigenmapping, a non-linear dimensionality reduction method, on the intrinsic functional connectivity data to find latent dimensions in the data. For each participant, we identified the latent dimension that was most similar to the published atlas of the principal gradient [41]. We then averaged the latent dimension across individuals to obtain regional values. The principal gradient in our data was highly correlated with the previously published atlas (r=0.90) as well as to the average anterior-to-posterior anatomical coordinates (r=0.91).

A.8 Representational capacity and Neurosynth

Neurosynth is an automated meta-analysis tool that uses natural language processing to search for the co-occurrence of behavioral terms and fMRI activation in a standardized coordinate form across over 10,000 studies [30]. The tool statistically links cognitive processes and mental states to regional brain activity. For replicability, we used previously described methods, the published principal gradient, and the version of Neurosynth accessed at the time of the previous study to create masks for the regions of interest [41]. Briefly, the published atlas was projected to a 2-mm volumetric space for our atlas (available here). The volumetric map was binned into five-percentile increments and binarized. Each of these 20 maps ranging from 0–5% to 95–100% were used as inputs to the meta-analysis. For each map, the output of the meta-analysis was a z-statistic associated with the feature term. The feature terms were derived from the set of 50 topic terms (as in previous work. This previous work removed 30 terms because they were above the statistical association threshold of z>3.1, removed 6 as presumed noise terms because they were not consistent with any particular cognitive function. The final set consisted of 24 distinct classes of behavior. For visualization, the terms were then ordered according to the weighted mean of the feature association.

A.9 RNNs and task variance

Following [31], we reproduced figures showing that the same RNN can effectively perform multiple working memory tasks following training (**Figure 5**). In this work, we analyzed the recurrent matrix, and for replicability we used the 20 trained networks from the prior study. The recurrent matrix consisted of 256 units with positive (excitatory) and negative (inhibitory) connections; across all 20 networks, these connection strengths range from -4.27 to 2.99. Noisy input units encoded task variables and output motor units read out nonlinearly to encode actions, such as a saccade or reach direction. A supervised learning protocol was used to modify all connection weights (input, recurrent, and output) to minimize the difference between the output units and the correct task output.

To match the range of values in the connectivity matrices extracted from human brain networks, we scaled the values of the recurrent connection matrix to between 0 and 1. When modeling random walks on these networks, this scaling also serves to bias the behavior of random walkers that represent activity flow to mimic inhibitory signaling. Specifically, connections that were originally more negative become lower probability nodes which inhibit the subsequent propagation to neighboring pathways and the overall flow of activity to inhibited units. Inhibitory pathways thus decrease the number of random walkers and redirect them away from locally connected, similar circuits (or clusters) of units towards different excitatory pathways.

The RNNs were trained to perform 20 tasks. Tasks included the Go, Reaction-time go (RT Go), Delayed go (Dly Go), Anti-response (Anti), Reaction-time anti-response (RT Anti), Delayed anti-response (Dly Anti), Decision making 1 (DM 1), Decision making 2 (DM 2), Context-dependent decision making 1 (Ctx DM 1), Context-dependent decision making 2 (Ctx DM 2), Multi-sensory decision making (MultSen DM), Delayed decision making 1 (Dly DM 1), Delayed decision making 2 (Dly DM 2), Context-dependent delayed decision making 1 (Ctx Dly DM 1), Context-dependent delayed decision making 2 (Ctx Dly DM 2), Multi-sensory delayed decision making (MultSen Dly DM), Delayed match-to-sample (DMS), Delayed non-match-to-sample (DNMS), Delayed match-to-category (DMC), and Delayed non-match-to-category (DNMC). These tasks were simplified from tasks used in prior non-human animal experiments and computational studies [66, 67, 68, 69, 70, 71, 72, 73]. Collectively, these tasks involve cognitive processes including pro-response, working memory, anti-response, decision making, gating modulation, integrating modulation, comparison, and categorization.

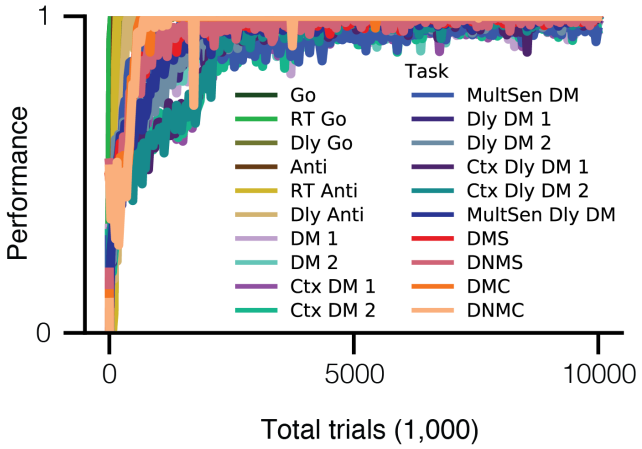

Figure 5: **Trained RNNs effectively perform working memory tasks.** Task performance increases with training. Reproduced with code from [31].

The task variance $\mathrm{TV}_i(A)$ for task A and unit i was calculated over many trials for each RNN unit. The averaged variance across all trials is the task variance:

$$\mathrm{TV}_i(A) = \left\langle \left[f_i(j,t) - \langle f_i\left(j',t\right)\rangle_{j'} \right]^2 \right\rangle_{j,t} \tag{4}$$

where $f_i(j,t)$ is the activity (non-negative and non-saturating firing rate) of unit i on time t of trial j.

A.10 Statistical analysis

We used generalized additive models with penalized splines, a non-parametric regression method, allowing for statistically rigorous modeling of linear and non-linear effects while minimizing overfitting [74]. To predict dimensionality we included covariates for the number of random walkers, the average degree of connectivity, window size, and splines for the principal gradient ($k = 4$) and its interaction with window size. To assess the relationship between the number of random walkers and the representational capacity, we used a Pearson correlation coefficient.

References

[1] David Attwell and Simon B Laughlin. An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, 21(10):1133–1145, 2001.

[2] Horace B Barlow et al. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1(01), 1961.

[3] Simon B Laughlin. Energy as a constraint on the coding and processing of sensory information. *Current opinion in neurobiology*, 11(4):475–480, 2001.

[4] Claude E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec*, 1959.

[5] Chris Eliasmith, Terrence C Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A large-scale model of the functioning brain. *science*, 338(6111): 1202–1205, 2012.

[6] Chris R Sims. Efficient coding explains the universal law of generalization in human perception. *Science*, 360(6389):652–656, 2018.

[7] Michael L Mack, Alison R Preston, and Bradley C Love. Ventromedial prefrontal cortex compression during concept learning. *Nature communications*, 11(1):1–11, 2020.

[8] Timothy F Brady, Talia Konkle, and George A Alvarez. Compression in visual working memory: using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, 138(4):487, 2009.

[9] Christopher J Bates, Rachel A Lerch, Chris R Sims, and Robert A Jacobs. Adaptive allocation of human visual working memory capacity during statistical and categorical learning. *Journal of vision*, 19(2):11–11, 2019.

[10] Shekoofeh Hedayati, Ryan E O'Donnell, and Brad Wyble. A model of working memory for latent representations. *Nature Human Behaviour*, 6(5):709–719, 2022.

[11] Lisa Feldman Barrett. *How emotions are made: The secret life of the brain*. Pan Macmillan, 2017.

[12] Lucy Lai and Samuel J Gershman. Policy compression: An information bottleneck in action selection. *Psychology of Learning and Motivation*, 74:195–232, 2021.

[13] Mandyam Veerambudi Srinivasan, Simon Barry Laughlin, and Andreas Dubs. Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 216(1205):427–459, 1982.

[14] Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. 2016.

[15] Matthew Farrell, Stefano Recanatesi, Timothy Moore, Guillaume Lajoie, and Eric Shea-Brown. Gradient-based learning drives robust representations in recurrent neural networks by balancing compression and expansion. *Nature Machine Intelligence*, pages 1–10, 2022.

[16] Matthew Chalk, Olivier Marre, and Gašper Tkačik. Toward a unified theory of efficient, predictive, and sparse coding. *Proceedings of the National Academy of Sciences*, 115(1): 186–191, 2018.

[17] Horace Barlow. Redundancy reduction revisited. *Network: computation in neural systems*, 12 (3):241, 2001.

[18] Andrea Avena-Koenigsberger, Bratislav Misic, and Olaf Sporns. Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, 19(1):17–33, 2018.

[19] Joaquín Goñi, Martijn P Van Den Heuvel, Andrea Avena-Koenigsberger, Nieves Velez de Mendizabal, Richard F Betzel, Alessandra Griffa, Patric Hagmann, Bernat Corominas-Murtra, Jean-Philippe Thiran, and Olaf Sporns. Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences*, 111(2):833–838, 2014.

[20] Caio Seguin, Adeel Razi, and Andrew Zalesky. Inferring neural signalling directionality from undirected structural connectomes. *Nature Communications*, 10, 2019.

[21] Richard Betzel, Joshua Faskowitz, Bratislav Misic, Olaf Sporns, and Caio Seguin. Multi-policy models of interregional communication in the human connectome. *bioRxiv*, 2022.

[22] Michael X. Henderson, Eli J. Cornblath, Adam S. Darwich, Bin Zhang, Hannah J Brown, Ronald J. Gathagan, Raizel M Sandler, Danielle S. Bassett, John Q. Trojanowski, and Virginia M. Y. Lee. α-synuclein pathology spread through the brain connectome is modulated by selective vulnerability and predicted by network analysis. *Nature neuroscience*, 22:1248 – 1257, 2019.

[23] Alessandra Griffa, Mathieu Mach, Julien Dedelley, Daniel Gutierrez-Barragan, Alessandro Gozzi, Gilles Allali, Joanes Grandjean, Dimitri Van De Ville, and Enrico Amico. The evolution of information transmission in mammalian brain networks. *bioRxiv*, 2022.

[24] Dale Zhou, Christopher W Lynn, Zaixu Cui, Rastko Ciric, Graham L Baum, Tyler M Moore, David R Roalf, John A Detre, Ruben C Gur, Raquel E Gur, et al. Efficient coding in the economics of human brain connectomics. *Network Neuroscience*, 6(1):234–274, 2022.

[25] Julia M. Huntenburg, Pierre-Louis Bazin, and Daniel S. Margulies. Large-scale gradients in human cortical organization. *Trends in Cognitive Sciences*, 22:21–31, 2018.

[26] Lilach Avitan and Carsen Stringer. Not so spontaneous: Multi-dimensional representations of behaviors and context in sensory areas. *Neuron*, 2022.

[27] Mark D. Humphries. Strong and weak principles of neural dimension reduction. *arXiv: Neurons and Cognition*, 2020.

[28] Stefano Fusi, Earl K Miller, and Mattia Rigotti. Why neurons mix: high dimensionality for higher cognition. *Current opinion in neurobiology*, 37:66–74, 2016.

[29] SueYeon Chung and LF Abbott. Neural population geometry: An approach for understanding biological and artificial neural networks. *arXiv preprint arXiv:2104.07059*, 2021.

[30] Tal Yarkoni, Russell A Poldrack, Thomas E Nichols, David C Van Essen, and Tor D Wager. Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8(8): 665–670, 2011.

[31] Guangyu Robert Yang, Madhura R. Joglekar, H. Francis Song, William T. Newsome, and Xiao-Jing Wang. Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, 22:297–306, 2019.

[32] Theodore D Satterthwaite, Mark A Elliott, Kosha Ruparel, James Loughead, Karthik Prabhakaran, Monica E Calkins, Ryan Hopson, Chad Jackson, Jack Keefe, Marisa Riley, et al. Neuroimaging of the philadelphia neurodevelopmental cohort. *Neuroimage*, 86:544–553, 2014.

[33] Fang-Cheng Yeh, Timothy D Verstynen, Yibao Wang, Juan C Fernández-Miranda, and Wen-Yih Isaac Tseng. Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PloS one*, 8(11):e80713, 2013.

[34] Matthew F Glasser, Timothy S Coalson, Emma C Robinson, Carl D Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F Beckmann, Mark Jenkinson, et al. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, 2016.

[35] Thomas M Cover and Joy A Thomas. Information theory and statistics. *Elements of information theory*, 1(1):279–335, 1991.

[36] Joaquín Goñi, Andrea Avena-Koenigsberger, Nieves Velez de Mendizabal, Martijn P. van den Heuvel, Richard F. Betzel, and Olaf Sporns. Exploring the morphospace of communication efficiency in complex networks. *PLOS ONE*, 8, 03 2013. doi: 10.1371/journal.pone.0058070. URL `https://doi.org/10.1371/journal.pone.0058070`.

[37] A Fornito, A Zalesky, and ET Bullmore. Chapter 7-paths, diffusion, and navigation. *Fundamentals of brain network analysis*, pages 207–255, 2016.

[38] Sarah E Marzen and Simon DeDeo. The evolution of lossy compression. *Journal of The Royal Society Interface*, 14(130):20170166, 2017.

[39] Peiran Gao, Eric Trautmann, Byron Yu, Gopal Santhanam, Stephen Ryu, Krishna Shenoy, and Surya Ganguli. A theory of multineuronal dimensionality, dynamics and measurement. *BioRxiv*, page 214262, 2017.

[40] Ashok Litwin-Kumar, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and LF Abbott. Optimal degrees of synaptic connectivity. *Neuron*, 93(5):1153–1164, 2017.

[41] Daniel S Margulies, Satrajit S Ghosh, Alexandros Goulas, Marcel Falkiewicz, Julia M Huntenburg, Georg Langs, Gleb Bezgin, Simon B Eickhoff, F Xavier Castellanos, Michael Petrides, et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, 113(44):12574–12579, 2016.

[42] Surya Ganguli and Haim Sompolinsky. Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Annual review of neuroscience*, 35:485–508, 2012.

[43] Evelyn Tang, Marcelo G Mattar, Chad Giusti, David M Lydon-Staley, Sharon L Thompson-Schill, and Danielle S Bassett. Effective learning is accompanied by high-dimensional and efficient representations of neural activity. *Nature neuroscience*, 22(6):1000–1009, 2019.

[44] Valerie J Sydnor, Bart Larsen, Danielle S Bassett, Aaron Alexander-Bloch, Damien A Fair, Conor Liston, Allyson P Mackey, Michael P Milham, Adam Pines, David R Roalf, et al. Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron*, 109(18):2820–2846, 2021.

[45] Mehrdad Jazayeri and Srdjan Ostojic. Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Current Opinion in Neurobiology*, 2021.

[46] Michael W. Cole, Jeremy R. Reynolds, Jonathan D. Power, Grega Repov, Alan Anticevic, and Todd Samuel Braver. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature neuroscience*, 16:1348 – 1355, 2013.

[47] Takuya Ito, Guangyu Robert Yang, Patryk Laurent, Douglas H Schultz, and Michael W Cole. Constructing neural network models from brain data reveals representational transformations linked to adaptive behavior. *Nature Communications*, 13(1):1–16, 2022.

[48] Yi Ma, Doris Tsao, and Heung-Yeung Shum. On the principles of parsimony and self-consistency for the emergence of intelligence. 2022. doi: 10.48550/ARXIV.2207.04630. URL `https://arxiv.org/abs/2207.04630`.

[49] John D. Murray, Alberto Bernacchia, David J. Freedman, Ranulfo Romo, Jonathan D. Wallis, Xinying Cai, Camillo Padoa-Schioppa, Tatiana Pasternak, Hyojung Seo, Daeyeol Lee, and Xiao-Jing Wang. A hierarchy of intrinsic timescales across primate cortex. *Nature neuroscience*, 17:1661 – 1663, 2014.

[50] Shi Gu, Fabio Pasqualetti, Matthew Cieslak, Qawi K Telesford, B Yu Alfred, Ari E Kahn, John D Medaglia, Jean M Vettel, Michael B Miller, Scott T Grafton, et al. Controllability of structural brain networks. *Nature communications*, 6(1):1–10, 2015.

[51] Andy Clark. Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, 36(3):181–204, 2013.

[52] Kevin D Himberger, Hsiang-Yun Chien, and Christopher J Honey. Principles of temporal processing across the cortical hierarchy. *Neuroscience*, 389:161–174, 2018.

[53] Jae Dong Noh and Heiko Rieger. Random walks on complex networks. *Physical review letters*, 92(11):118701, 2004.

[54] Rich Pang, Benjamin James Lansdell, and Adrienne L. Fairhall. Dimensionality reduction in neuroscience. *Current Biology*, 26:R656–R660, 2016.

[55] Stephen J Gotts, Adrian W Gilmore, and Alex Martin. Brain networks, dimensionality, and global signal averaging in resting-state fmri: Hierarchical network structure results in low-dimensional spatiotemporal dynamics. *NeuroImage*, 205:116289, 2020.

[56] David R Roalf, Megan Quarmley, Mark A Elliott, Theodore D Satterthwaite, Simon N Vandekar, Kosha Ruparel, Efstathios D Gennatas, Monica E Calkins, Tyler M Moore, Ryan Hopson, et al. The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. *Neuroimage*, 125:903–919, 2016.

[57] Marcus E Raichle and Mark A Mintun. Brain work and brain imaging. *Annu. Rev. Neurosci.*, 29:449–476, 2006.

[58] Clotilde Lecrux, Miled Bourourou, and Edith Hamel. How reliable is cerebral blood flow to map changes in neuronal activity? *Autonomic Neuroscience*, 217:71–79, 2019.

[59] Ruben C Gur, J Daniel Ragland, Martin Reivich, Joel H Greenberg, Abass Alavi, and Raquel E Gur. Regional differences in the coupling between resting cerebral blood flow and metabolism may indicate action preparedness as a default state. *Cerebral Cortex*, 19(2):375–382, 2009.

[60] S Neil Vaishnavi, Andrei G Vlassenko, Melissa M Rundle, Abraham Z Snyder, Mark A Mintun, and Marcus E Raichle. Regional aerobic glycolysis in the human brain. *Proceedings of the National Academy of Sciences*, 107(41):17757–17762, 2010.

[61] Ze Wang, Geoffrey K Aguirre, Hengyi Rao, Jiongjiong Wang, María A Fernández-Seara, Anna R Childress, and John A Detre. Empirical optimization of asl data analysis using an asl data processing toolbox: Asltbx. *Magnetic resonance imaging*, 26(2):261–269, 2008.

[62] Theodore D Satterthwaite, Russell T Shinohara, Daniel H Wolf, Ryan D Hopson, Mark A Elliott, Simon N Vandekar, Kosha Ruparel, Monica E Calkins, David R Roalf, Efstathios D Gennatas, et al. Impact of puberty on the evolution of cerebral perfusion during adolescence. *Proceedings of the National Academy of Sciences*, 111(23):8643–8648, 2014.

[63] Wen-Chau Wu, Varsha Jain, Cheng Li, Mariel Giannetta, Hallam Hurt, Felix W Wehrli, and Danny JJ Wang. In vivo venous blood t1 measurement using inversion recovery true-fisp in children and adults. *Magnetic resonance in medicine*, 64(4):1140–1147, 2010.

[64] Varsha Jain, Jeffrey Duda, Brian Avants, Mariel Giannetta, Sharon X Xie, Timothy Roberts, John A Detre, Hallam Hurt, Felix W Wehrli, and Danny JJ Wang. Longitudinal reproducibility and accuracy of pseudo-continuous arterial spin–labeled perfusion mr imaging in typically developing children. *Radiology*, 263(2):527–536, 2012.

[65] Reinder Vos de Wael, Oualid Benkarim, Casey Paquola, Sara Lariviere, Jessica Royer, Shahin Tavakol, Ting Xu, Seok-Jun Hong, Georg Langs, Sofie Valk, et al. Brainspace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Communications biology*, 3(1):1–10, 2020.

[66] Shintaro Funahashi, Charles J Bruce, and Patricia S Goldman-Rakic. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of neurophysiology*, 61(2): 331–349, 1989.

[67] Douglas P Munoz and Stefan Everling. Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, 5(3):218–228, 2004.

[68] Joshua I Gold, Michael N Shadlen, et al. The neural basis of decision making. *Annual review of neuroscience*, 30(1):535–574, 2007.

[69] Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474):78–84, 2013.

[70] David Raposo, Matthew T Kaufman, and Anne K Churchland. A category-free neural population supports evolving demands during decision-making. *Nature neuroscience*, 17(12):1784–1792, 2014.

[71] Ranulfo Romo, Carlos D Brody, Adrián Hernández, and Luis Lemus. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, 399(6735):470–473, 1999.

[72] Earl K Miller, Cynthia A Erickson, and Robert Desimone. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of neuroscience*, 16(16):5154–5167, 1996.

[73] David J Freedman and John A Assad. Neuronal mechanisms of visual categorization: an abstract view on decision making. *Annual review of neuroscience*, 39:129–147, 2016.

[74] Simon N Wood. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99(467):673–686, 2004.