



Optimizing the human learnability of abstract network representations

William Qian^a, Christopher W. Lynn^{b,c,d}, Andrei A. Klishin^e, Jennifer Stiso^e, Nicolas H. Christianson^f, and Dani S. Bassett^{a,e,g,h,i,j,1}

Edited by Marcus Raichle, Mallinckrodt Institute of Radiology and Department of Neurology, Washington University of School of Medicine, St. Louis, MO; received November 23, 2021; accepted June 30, 2022

Precisely how humans process relational patterns of information in knowledge, language, music, and society is not well understood. Prior work in the field of statistical learning has demonstrated that humans process such information by building internal models of the underlying network structure. However, these mental maps are often inaccurate due to limitations in human information processing. The existence of such limitations raises clear questions: Given a target network that one wishes for a human to learn, what network should one present to the human? Should one simply present the target network as-is, or should one emphasize certain parts of the network to proactively mitigate expected errors in learning? To investigate these questions, we study the optimization of network learnability in a computational model of human learning. Evaluating an array of synthetic and real-world networks, we find that learnability is enhanced by reinforcing connections within modules or clusters. In contrast, when networks contain significant core-periphery structure, we find that learnability is best optimized by reinforcing peripheral edges between low-degree nodes. Overall, our findings suggest that the accuracy of human network learning can be systematically enhanced by targeted emphasis and de-emphasis of prescribed sectors of information.

graph learning | maximum entropy | complex networks

From a young age, humans demonstrate the capacity to learn the relationships between concepts (1–3). During the learning process, humans are exposed to discrete chunks of information that combine and interconnect to form cognitive maps that can be represented as complex networks (4–9). These chunks of information often appear in a natural sequential order, such as words in language, notes in music, and abstract concepts in stories and classroom lectures (10–14). Further, these sequences are encoded in the brain as networks, with links between items reflecting observed transitions (see refs. 15–18 for empirical studies and ref. 19 for a recent review). Broadly, the fact that many different types of information exhibit temporal order (and therefore network structure) motivates investigations into the processes that underlie the human learning of transition networks (8, 19, 20).

To understand the network-learning process, recent studies have investigated how humans internally construct abstract representations of associations (21–23). Using a variety of approaches, from computational models to artificial neural networks, such studies have consistently found that the mind builds network representations by integrating information over time. Such integration enables humans to compress exact sequences of experienced events into broader, but less precise, representations of context (24). These mental representations allow learners to make better generalizations about new information, at the cost of accuracy (22). Here, we focus on one particular modeling approach that accounts for the temporal integration and inaccuracies inherent in human learning. In particular, we build upon a maximum-entropy model, which posits that the mind learns a network representation of the world in a manner guided by a tradeoff between accuracy and complexity (21, 25). Specifically, in order to conserve mental resources, humans will tend to reduce the complexity of their representations at the cost of accuracy by allowing for errors during the learning process.

While inaccuracies in human learning can aid flexibility across contexts, they present fundamental obstacles for the human comprehension of transition networks. Thus, a clear question emerges: What strategies should be employed to most effectively communicate the structure of a network to an inaccurate human learner? Prior studies of animal communication and behavior have demonstrated the utility of exaggerating the presentation of certain signals to receivers in offsetting erroneous information processing (26, 27). Similarly, one could imagine that, by emphasizing some features of a network over others, one may be able to correct for errors in human learning. Such an approach of targeted modulation of emphasis may be helpful not only in learning a whole network,

Significance

Information can often be viewed as a network of associations between concepts. Humans build mental models of information networks in the world around them, yet those models consistently contain some errors. Here, we present a computational framework for simulating the optimization of human network learning by intentionally emphasizing or exaggerating some network features over others. We demonstrate in a computational model of human learning that targeted emphasis and de-emphasis can substantially enhance a learner's grasp of network structure. Further, we identify how optimal emphasis patterns vary with the topology of the target network structure to be learned, as well as the baseline accuracy of the human learner. Our findings illuminate the principles of design and the optimization of network learnability.

Author contributions: W.Q. and D.S.B. designed research; W.Q. performed research; C.W.L., A.A.K., and N.H.C. contributed new reagents/analytic tools; W.Q. analyzed data; and W.Q., J.S., and D.S.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: dsb@seas.upenn.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2121338119/-DCSupplemental>.

Published August 22, 2022.

but also in optimally learning particularly challenging parts of a network. In fact, humans show consistent difficulties in learning certain motifs in networks, such as the connections between modules (21, 28–30). Taken together, these observations suggest that disproportionately weighting specific network features that are difficult to learn may facilitate human network learning.

Mathematical Methods

To study the optimization of network learnability, we first require a model describing how humans learn networks. Following ref. 21, we write down an analytic model that captures a wide range of human behaviors observed in network-learning experiments. Specifically, we consider a transition probability matrix $A \in \mathbb{R}^{n \times n}$ that describes random walks on the observed network, where A_{ij} represents the probability of a transition to node j when starting from node i . Then, after observing a sufficient number of random walks, the human’s internal representation of the network converges to the following analytic form:

$$f(A) = (1 - e^{-\beta})A(I - e^{-\beta}A)^{-1}, \quad [1]$$

where the single parameter β reflects the accuracy of the human learner. In the limit $\beta \rightarrow 0$, the learned network structure $f(A)$ is a fully connected network with uniform edge weights and, hence, bears no resemblance to the actual network A . In contrast, in the limit $\beta \rightarrow \infty$, the learning process is free of errors, and the learned network structure $f(A)$ is an exact replica of the actual network A .

The following question then arises: Given a transition network A and a human learner with accuracy β , what is the optimal input transition network $A_{\text{in}} = A^*$ that, when presented to a human learner, results in a perceived network structure $f(A_{\text{in}})$ that most closely matches the true structure A ? In general, there is no reason to presume that it is optimal to present the learner with the true network (such that $A^* = A$). Indeed, teaching and other forms of communication often involve the purposeful emphasis or exaggeration of some pieces of information over others (31–33). Thus, it is possible that modulating emphasis on certain network features, in a precise and targeted manner that serves to counteract natural biases or expected errors, might enhance learnability. Moreover, optimal emphasis strategies are likely to vary from person to person, depending critically upon the accuracy β of the human’s learning process.

One natural approach to answering this generic question is to find the input matrix A_{in} such that the learned representation $f(A_{\text{in}})$ is equal to the target network A . From Eq. 1, one can derive for a given β an analytical form for the input $A_{\text{in}} = A^*$ such that $f(A^*) = A$ holds exactly. While this approach is mathematically elegant, it has limited application to real-world scenarios because the resulting A^* often contains negative entries or requires inverting a singular matrix and is, hence, an ill-defined transition matrix (SI Appendix, Fig. S1). To overcome this hurdle, one can instead characterize how a perceived network structure $f(A_{\text{in}})$ diverges from some true structure A using the Kullback–Leibler divergence $D_{KL}(A||f(A_{\text{in}}))$. We choose the Kullback–Leibler divergence for our analyses because of its connections with information theory and its usage in previous papers that study this model of human network learning (20, 21, 25). In particular, the Kullback–Leibler divergence, as used in this context, can be interpreted as a measure of the inefficiency of the learned representation. To determine the optimal input $A_{\text{in}} = A^*$ such that A^* is a well-defined transition matrix, we determine a weighted network with adjacency matrix G_{in} such that the corresponding

transition matrix A_{in} minimizes the Kullback–Leibler divergence $D_{KL}(A||f(A_{\text{in}}))$ between the learned structure $f(A_{\text{in}})$ and the true structure A . Practically, we implement this strategy using dual annealing, which is a powerful and common method for bounded optimization (34–36). For simplicity, we restrict our analysis to undirected networks G_{in} .

For large input networks with many edges, this optimization process can become computationally unwieldy. To address this issue, we only consider input networks G_{in} that respect the symmetries of A : All structurally unique edges in A must have the same edge weight in G_{in} . Further, we only consider the inclusion of edges in G_{in} when they have a counterpart in A with nonzero weight. In this manner, the network-optimization process can be parameterized by a significantly smaller number of trainable values for networks with a high degree of symmetry.

To investigate possible strategies for enhancing transition-network learnability, we apply the optimization method to two transition networks: a modular network and a lattice network. Both of these networks have 15 vertices and share the property that every node has degree 4. Importantly, previous human experiments were able to directly estimate the accuracy parameter β of human learning in these networks (21). Next, to explore the optimization of learnability in asymmetric networks with nonuniform transition probabilities, we consider the optimization of learnability for networks constructed from generative network models. Lastly, to probe how real-world information networks ought to be designed, we investigate how learnability can be maximized for semantic networks extracted from college mathematics textbooks.

In performing these numerical experiments, we are guided by several hypotheses. Specifically, in considering prior work demonstrating the efficacy of exaggeration in animal communication (26, 27), we predict that strategic modulation of emphasis will significantly improve learnability in both synthetic and real-world information networks. Furthermore, in view of prior work demonstrating that highly clustered networks are more learnable than lattice or random networks (25), we hypothesize that optimal emphasis-modulation strategies will reinforce connections within clusters and de-emphasize connections across clusters. And, finally, given the human tendency to attend to salient information, we hypothesize that core–periphery networks will be best learned by emphasizing edges in the network periphery, which would otherwise be less easily learned. Taken together, our numerical experiments aim to assess whether and how the accuracy of human network learning can be systematically enhanced by targeted emphasis and de-emphasis of network features.

Results

Optimizing the Learnability of Graph Exemplars.

The modular graph exemplar. We begin by studying the learning optimization of the modular graph shown in Fig. 1A, which has been used in human learning studies (15, 28–30). In this graph, there are only three structurally unique edges: cross-cluster edges (orange), boundary edges that are adjacent to cross-cluster edges (green), and edges deep within modules (gray). Thus, the structure of A_{in} can be determined by two free parameters, λ_{cc} and λ_b , representing the weights of cross-cluster and boundary edges in G_{in} , respectively, relative to the weight of deep edges in G_{in} . Note that one free parameter has been removed due to the constraint that A_{in} is normalized, such that all rows sum to one.

To illustrate the parameter regimes where certain weighting combinations of λ_{cc} and λ_b are effective in enhancing learnability,

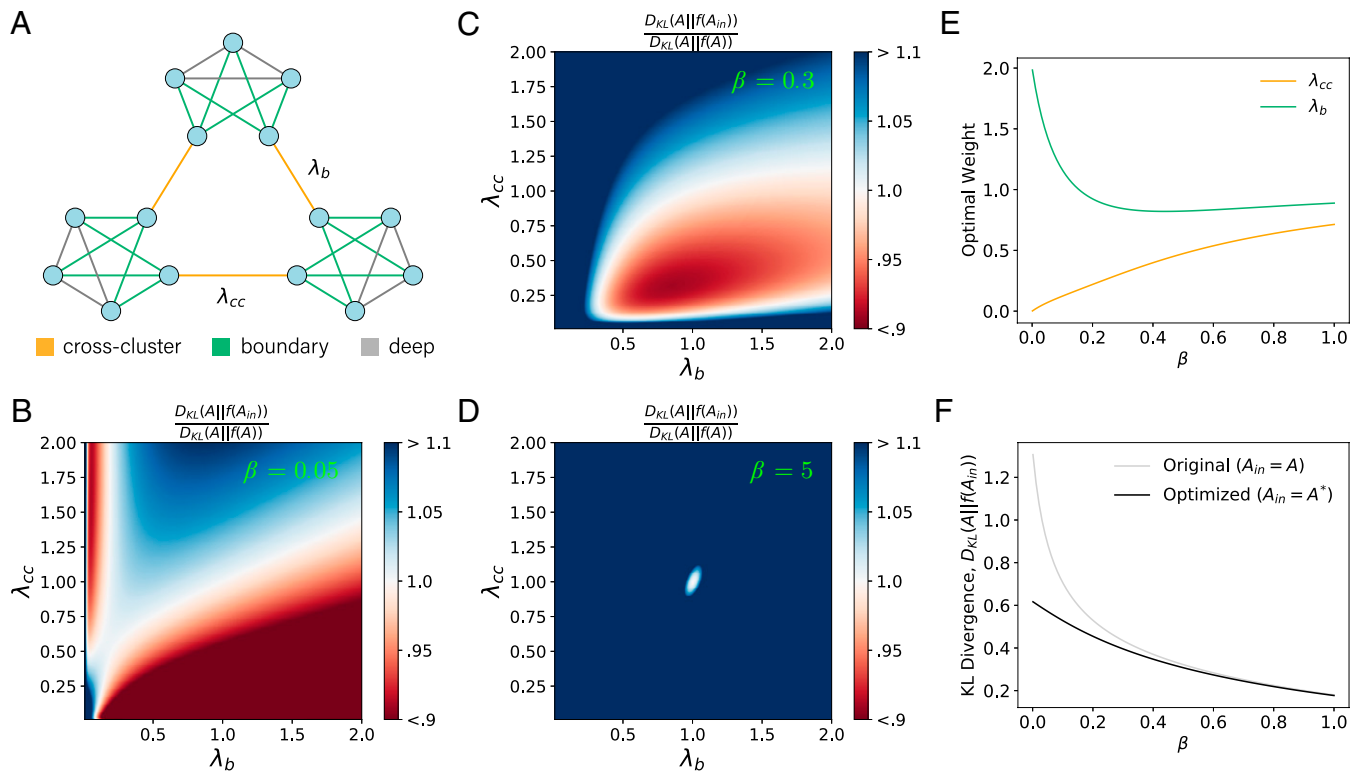


Fig. 1. Optimizing the learnability of a modular graph. (A) A modular graph with 15 nodes, each with degree $k_i = 4$, resulting in 30 edges. (B–D) Here, we show the Kullback–Leibler divergence ratio (less than one indicates enhanced learnability) across a section of the λ_{cc} , λ_b parameter space, for different values of β . For increased contrast, the ratios have been truncated to the range [0.9, 1.1]. (B) Results for $\beta = 0.05$, corresponding roughly to the median accuracy of human learners in prior studies (21). (C) Results for $\beta = 0.3$, corresponding to the mean accuracy of human learners in prior studies (21). (D) Results for $\beta = 5$, corresponding to an exceptionally accurate network learner. (E) The optimal edge weights λ_{cc} and λ_b for $0 < \beta < 1$. (F) The Kullback–Leibler divergence between the learned network and the true network for different values of β , both with and without input network optimization.

we first computed the ratio $\frac{D_{KL}(A||f(A_{in}))}{D_{KL}(A||f(A))}$ over the parameter space $0 \leq \lambda_{cc} \leq 2$, $0 \leq \lambda_b \leq 2$, for three different values of β (Fig. 1 B–D). This ratio characterizes the learnability that can be achieved by targeted modulation of emphasis in the network. Specifically, a ratio of less than one would indicate that the emphasized network improves learnability over the true network.

Interestingly, at low values of β (Fig. 1B; $\beta = 0.05$), when the learning process is highly inaccurate, there are two regimes in which an emphasized network structure improves learnability: one that heavily de-emphasizes boundary edges and one that moderately de-emphasizes cross-cluster edges. For intermediate values of β (such as $\beta = 0.3$ in Fig. 1C), the two optimal regimes combine into one. As the learning accuracy increases further (Fig. 1D; $\beta = 5$), the one optimal regime decreases in size and converges to the true network structure. Thus, for extremely precise learners, the only reasonable network structure to learn would be the true network structure, corresponding to $\lambda_{cc} = \lambda_b = 1$ (Fig. 1D; $\beta = 5$).

To assess the precise values of edge weights that lead to optimal learning of the modular graph, we minimize $D_{KL}(A||f(A_{in}))$ with respect to λ_{cc} and λ_b at different values of β (Fig. 1E). We find that cross-cluster edges are always de-emphasized, whereas boundary edges are overemphasized for inaccurate learners, but de-emphasized for the average human learner ($\beta \approx 0.3$). We present a graphical depiction of the optimal input network and the resulting learned structure for the modular graph at $\beta = 0.05$ in Fig. 2A.

The lattice-graph exemplar. To understand how optimizing network learnability varies with the topology of the target network, we also study the optimization of learnability of a lattice graph that was examined in human learning studies (21, 29)

(SI Appendix, Fig. S2). While we find qualitative differences in the efficacy of the optimization process for modular and lattice graphs, we find, similarly, that small, cluster-like substructures in the lattice graph are emphasized to maximize learnability, whereas edges between these substructures are de-emphasized. We present a graphical depiction of the optimal input network and the resulting learned structure for the lattice graph at $\beta = 0.05$ in Fig. 2B.

A Sierpiński graph exemplar. Next, to assess whether the strategy of overemphasizing edges within clusters and de-emphasizing those between clusters extends to larger networks with more complex community organization, we also consider a Sierpiński network with hierarchical community structure (SI Appendix, Fig. S3). Consistent with previous findings, we find that de-emphasizing cross-cluster edges at all hierarchical levels of organization is an effective strategy for optimizing learnability. In addition, we observe that cross-cluster edges at the highest level of organization ought to be de-emphasized the most.

Optimizing the Learnability of Generated Networks.

Stochastic block networks. Thus far, we have found that the learnability of networks with modular structure is optimized by overemphasizing the edges within clusters of nodes and de-emphasizing the edges between clusters. However, our analyses have focused on networks with high degrees of structural symmetry. Here, we extend our analysis to randomly generated networks, studying the optimization of learnability in stochastic block networks.

We consider two classes of stochastic block networks: 1) stochastic block networks in the absence of specific structure–degree correlations, where all cross-cluster edges are equally

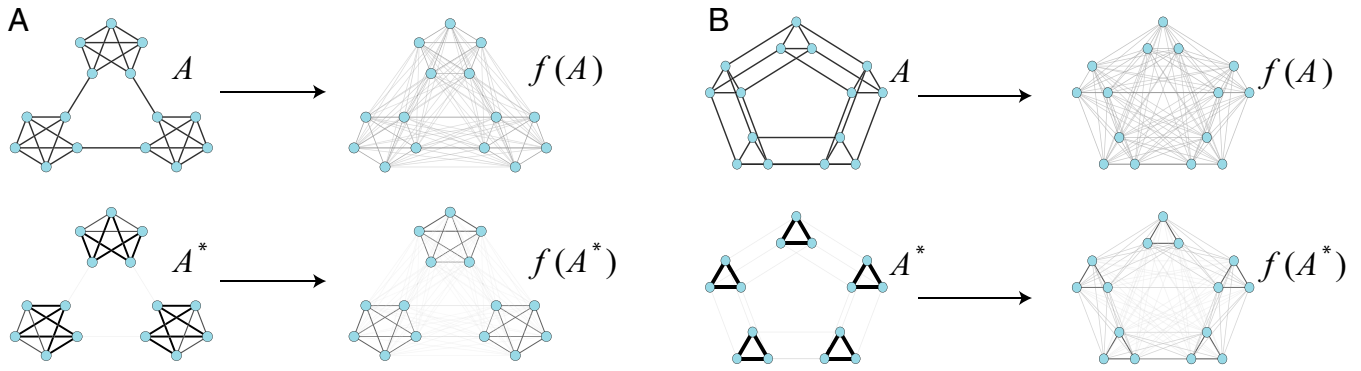


Fig. 2. Optimal emphasis modulation of the modular and lattice networks. Here, we show the learned networks resulting from human learning of the modular and lattice networks, respectively (*A* and *B*, *Upper*), as well as from the modular and lattice networks optimized for learnability (*A* and *B*, *Lower*). Optimized and learned networks were both computed at $\beta = 0.05$. Edge thickness indicates transition probabilities.

likely to be included, and all within-cluster edges are equally likely to be included; and 2) degree-corrected stochastic block networks with heterogeneous degree distributions. These classes were chosen to assess whether degree heterogeneity, a common feature of real-world networks (37, 38), influences the efficacy of strategies for enhancing network learnability in modular networks. In particular, we consider networks with $N = 200$ nodes, 5 communities, and an average degree of $\langle k \rangle = 10$ (Fig. 3 *A* and *B*; see *Materials and Methods* for network-generation procedures). For a given stochastic block network G_{SBM} with a normalized transition matrix A_{SBM} , we parameterize the network presented to learners G_{in} by a single parameter λ_{cc} , representing the weight of edges between clusters relative to the weight of

edges within clusters. We then compute the cross-cluster weight λ_{cc} that optimizes the learnability of the transition network A_{SBM} .

We begin by analyzing the enhancement of learnability for stochastic block networks without structure–degree correlations. In Fig. 3*C*, we show the optimal cross-cluster weight λ_{cc} as a function of the fraction f of edges chosen to be within communities. Importantly, we find that the optimal cross-cluster weight decreases considerably as the modularity of the target stochastic block networks increases. Moreover, for higher β , we find that optimally emphasized networks maintain more weight on cross-cluster edges, consistent with our earlier analysis of the modular network (Fig. 1*A*). In addition, increases in the learnability of

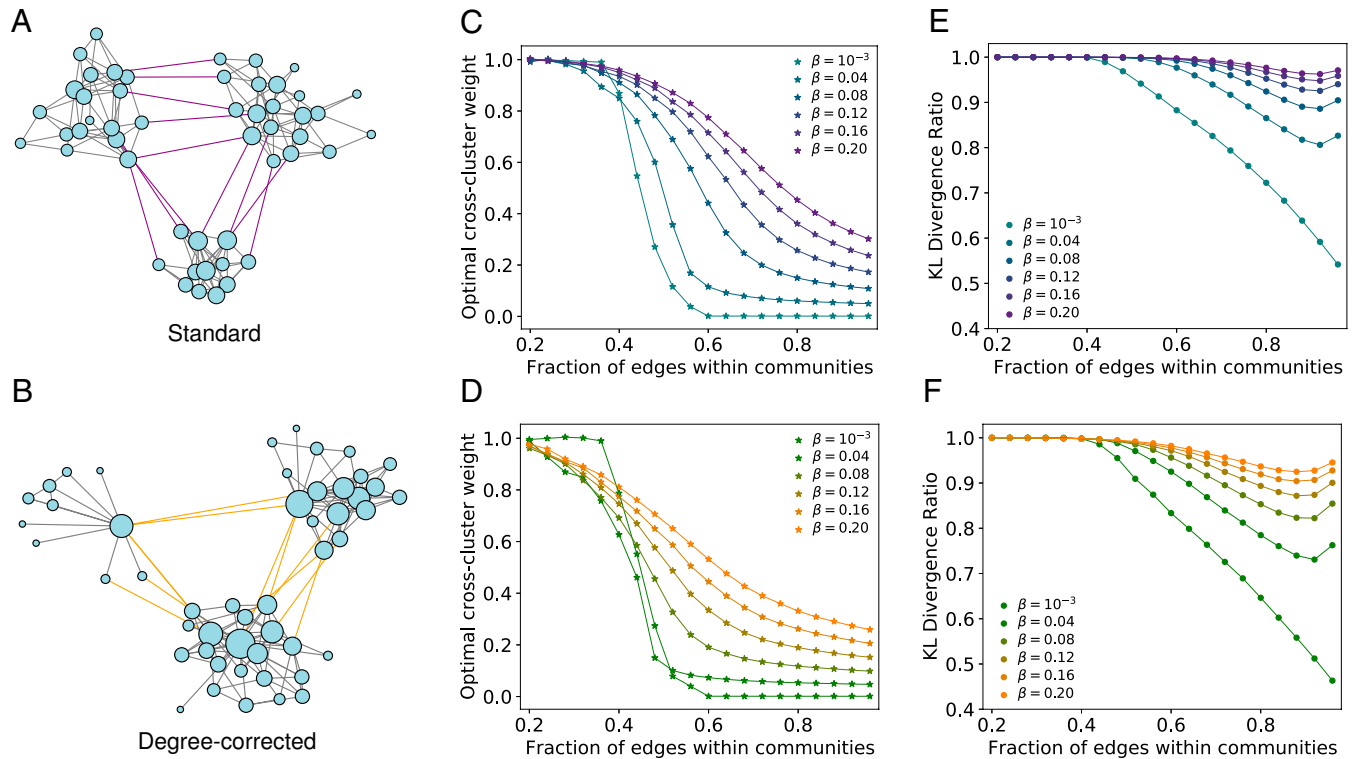


Fig. 3. Optimizing the learnability of synthetic modular networks. (*A* and *B*) Examples of a standard stochastic block network and a degree-corrected stochastic block network. Node sizes are proportional to node degrees, with cross-cluster edges shown in purple and orange, respectively. (*C* and *D*) The optimal cross-cluster edge weight λ_{cc} for enhancing learnability versus the fraction f of edges within communities at different values of β . Results are shown for stochastic block networks and degree-corrected stochastic block networks, respectively. (*E* and *F*) The Kullback–Leibler (KL) divergence ratio $\frac{D_{KL}(A|f(A_{in}))}{D_{KL}(A|f(A))}$ achieved with optimal cross-cluster edge weights at different values of β . Results are shown for stochastic block networks and degree-corrected stochastic block networks, respectively. The findings reported in *C–F* represent results obtained for networks with $N = 200$ nodes, 5 communities, and an average degree of $\langle k \rangle = 10$. Each curve is an average over the results from 25 generated networks.

stochastic block networks (reductions of Kullback–Leibler divergence ratios) are most prominent for values of f above 0.8 and increase considerably as β decreases (Fig. 3E).

To determine whether the degree heterogeneity impacts the optimization of network learnability, we study degree-corrected stochastic block networks. For such networks, we find that the optimal cross-cluster edge weight λ_{cc} decreases faster with increasing f than for standard stochastic block networks (Fig. 3D). Interestingly, for degree-corrected stochastic block networks, the improvements in learnability are significantly larger than for regular stochastic block networks (Fig. 3E and F). For both types of networks, increases in network learnability are most pronounced for low values of β and peak for highly clustered networks with a fraction of within-community edges $f = 0.92$. This finding indicates that highly modular networks are most optimizable through cross-cluster weight tuning.

Watts–Strogatz networks. Just as we generalized the analysis of modular networks to stochastic block networks, we can also extend our analysis of lattice networks to a wide range of randomly generated Watts–Strogatz networks (SI Appendix, Fig. S4). Consistent with previous analysis, we find that edges in Watts–Strogatz networks that contribute to local, lattice-like clustering are emphasized when maximizing learnability.

Optimizing the Learnability of Semantic Networks Extracted from Mathematics Textbooks

Our results thus far have demonstrated that, for many classes of synthetic networks, edges that contribute to local clustering or intramodular connections are reinforced to maximize learnability. Still, it remains to be demonstrated that these results extend more generally to real-world information networks. To probe the optimal emphasis modulation strategies of real-world networks, we study semantic networks extracted from college-level linear algebra textbooks (39, 40). Specifically, nodes represent recurring concepts (e.g., “vector space” and “invertible”), and edges between concepts are weighted by the number of sentences in which the two concepts co-occur.

In previous analyses, we were able to reduce the number of free parameters in the network-learnability optimization process by considering either network symmetry or a partitioning of the edges into classes that are made distinguishable by the network-generation process (e.g., cross-cluster edges or nonring edges). However, given that these semantic networks are empirical, when optimizing a network representation to maximize learnability, we cannot reduce the number of optimization parameters a priori, and, instead, we must vary all edges with nonzero weight as free parameters. Specifically, for some semantic network $G_{SEM} = (V, E)$ with edge weights w_{ij} and normalized transition matrix A_{SEM} , we determine a weighted graph $G_{in} = (V, E)$ with weights w_{ij}^* such that its corresponding normalized transition matrix A_{in} minimizes $D_{KL}(A_{SEM} || f(A_{in}))$. In particular, for every edge $e = \{i, j\} \in E$, the factor $\lambda_{ij} = \frac{w_{ij}^*}{w_{ij}}$ by which the edge e is scaled in G_{in} is used as an optimization parameter. As before, these parameters are then simultaneously optimized via dual annealing to minimize the Kullback–Leibler divergence cost function. Then, we reduce the number of free parameters by one by enforcing the requirement that the total sum of edge weights in G_{in} equals that of G_{SEM} . Doing so allows for more interpretable comparisons between the optimized network A^* and the original semantic network A_{SEM} .

Interestingly, we find that for these semantic networks, very little improvement in learnability can be achieved at extremely low values of β (≈ 0 , reflecting poor learning accuracy), but significant

enhancement of learnability is possible for all other values of β (Fig. 4C). In particular, the network-optimization process yields the most benefit for moderately accurate learners ($\beta \approx 0.05$). This observation contrasts greatly with our prior findings in studying modular networks: that the greatest improvements in learnability occur near $\beta \approx 0$, with significantly diminishing improvements at higher β values. One natural explanation for this difference is that the semantic networks do not possess large-scale community structure, but, rather, can be characterized as possessing core–periphery structure with community structure within the periphery nodes (39). Therefore, the learnability-optimization strategies for modular networks—which mainly involved overemphasizing edges within clusters and de-emphasizing cross-cluster edges—are likely not applicable to increasing learnability of these semantic networks.

To understand which edges are reinforced or de-emphasized to increase learnability, for each semantic network, we follow the procedure outlined in ref. 39 and classify nodes into core and periphery categories (*Materials and Methods*). Then, for each network, we determine the community structure of the periphery nodes and categorize the edges of each network into four categories (Fig. 4A): edges between core nodes; edges between a core node and a periphery node; cross-cluster edges between periphery nodes; and within-cluster edges between periphery nodes. For each class of edges, we compute the mean optimal weight scaling over all 10 semantic networks for $0 < \beta \leq 0.2$ (Fig. 4B). Notably, we find that for $\beta > 0.05$, edges between core nodes are de-emphasized the most, whereas edges between periphery nodes are reinforced. This finding is sensible, as the core of a core–periphery network is densely connected. Therefore, any particular edge within the core could be de-emphasized to suppress potential spurious connections to nearby periphery nodes resulting from inaccurate learning. In addition, among the two classes of periphery–periphery edges, those that connect two nodes within the same periphery community tend to be exaggerated, whereas cross-cluster periphery–periphery edges are de-emphasized. These observations regarding cross-cluster and within-cluster periphery–periphery edges are consistent with our prior analyses of optimal cross-cluster edge weights in modular networks, which also suggest that cross-cluster weights should be de-emphasized.

We can further characterize the types of edges that are either overemphasized or underemphasized by comparing changes in edge weight with structural measures of centrality (*Materials and Methods*). First, we consider the relationship between edge-weight scaling and edge-betweenness centrality, a metric that quantifies the frequency with which shortest paths pass through a given edge. Given that cross-cluster edges have high edge-betweenness, it is natural to expect that edges with high betweenness will be de-emphasized when optimizing learnability. Indeed, we observe that, aside from edges with an edge-betweenness centrality of zero, there is a clear inverse relationship between edge-betweenness centrality and optimal edge-weight scaling in semantic networks (Fig. 4E; $\beta = 0.2$).

We also consider edge-weight scaling and edge-degree centrality, a metric that quantifies the average weighted degree of the two connected nodes. Edges within the core of a core–periphery network are likely to be incident on nodes with greater connectivity and, thus, would generally have higher edge-degree centrality. Thus, we expect that edge-degree centrality will be inversely related to optimal edge-weight scaling. Consistent with these expectations, we observe that edges with lower edge-degree centrality tended to be reinforced more, and edges with higher edge-degree centrality tended to be de-emphasized more (Fig. 4F).

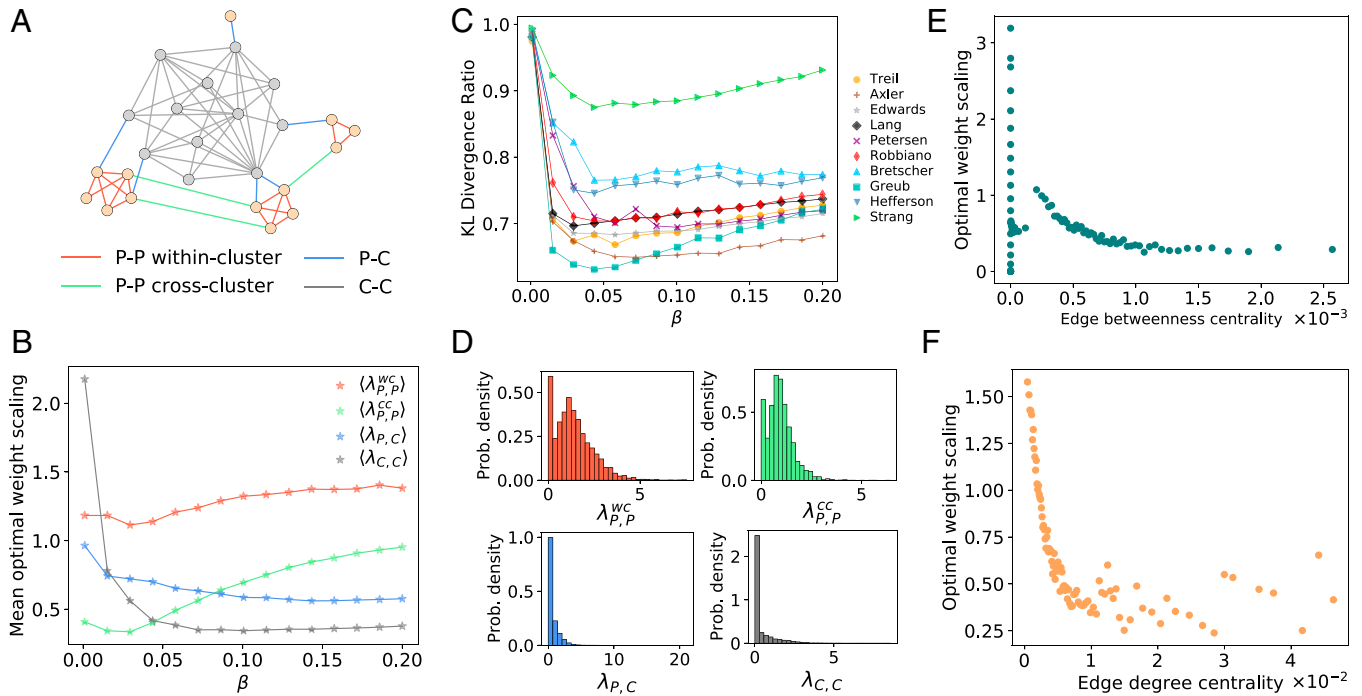


Fig. 4. Optimizing the learnability of semantic networks extracted from college mathematics textbooks. (A) A schematic of how edges in the semantic networks were classified based on core-periphery node classification and periphery community structure. (B) The optimal weight scaling for each of the four classes of edges shown at different values of β , averaged over all semantic networks. (C) The Kullback–Leibler (KL) divergence ratio $\frac{D_{KL}(A||f(A_{in}))}{D_{KL}(A||U(A))}$ achieved with optimized weight scaling at different values of β . Results are shown for each of the 10 semantic networks corresponding to the 10 college-level linear algebra textbooks (61–70). (D) The distribution of optimized edge-weight scalings for the classes of edges at $\beta = 0.2$, aggregated over all semantic networks. Prob., probability. (E and F) The optimal edge-weight scaling versus edge-betweenness centrality and edge-degree centrality, respectively, aggregated over all semantic networks for $\beta = 0.2$. Each datapoint represents an average over 500 edges binned by centrality score.

Performance of Network-Optimization Strategies in the Transient Network-Learning Regime.

Thus far, our analysis has implicitly assumed that human network learners are allowed to observe infinite sequences of network transitions (21, 25). While this assumption greatly simplifies optimization strategies, it is possible that such strategies break down when the number of transitions that human learners are allowed to observe is limited. To address this possibility, we ran simulations of the network-learning process in the transient regime with three different network-learning strategies (*Materials and Methods*): 1) maximum-likelihood estimation (optimal in the infinite observation limit), 2) standard human network learning (as reported in ref. 21), and 3) optimized human learning (as described here in our paper). Both the standard and optimized human learning strategies were evaluated at $\beta = 0.1$, which is close to the mean learning accuracy reported in ref. 21. These simulations were run for the modular graph with 15 nodes (Fig. 1A) and the semantic networks extracted from the linear algebra textbooks authored by Axler and Edwards (Fig. 1B and C, respectively).

For the modular network, optimized human learning maintained an edge in accuracy over standard human network learning, as the number of transitions observed increased (Fig. 5A). Both learning strategies were outperformed by maximum-likelihood estimation throughout the duration of the simulated learning processes. Remarkably, for the semantic networks, both human learning and optimized human learning initially outperformed the accuracy of maximum-likelihood estimation, with optimized human learning maintaining its superiority for a longer duration (Fig. 5B and C). One plausible explanation is that the inductive biases introduced in the human learning process enable humans to initially learn clustered areas of networks more efficiently than unbiased maximum-likelihood estimation strategies.

Discussion

In this article, we study how networks presented to human learners can be tuned to increase learnability. In particular, using a computational model of human learning, we compute the optimal network to present to a human learner so as to minimize the discrepancy between the learned representation and the target network. First, these methods were used to analyze two simple networks: a modular graph and a lattice graph. We find that for both graphs, improvements in learnability can be made by de-emphasizing edges that connect different modules or clusters of nodes and by exaggerating edges within modules or small clusters. This finding is consistent with studies of *in silico* models and *in vivo* animal behavior of sampling spaces. Animals and computational models exhibit nonrandom patterns of exploration in order to better sample an environment with nonregular network structure (41), effectively emphasizing and overrepresenting specific harder-to-learn portions of the environment (42). Further, these improvements increase considerably in magnitude for highly inaccurate human learners, but are less advantageous for accurate learners. Importantly, for inaccurate learners, the optimal input networks for both modular and lattice graphs result in internal network representations that capture clusters in the original network in a near-perfect manner, but poorly capture edges between small clusters or modules. Notably, edges between communities or clusters are already naturally difficult to learn in the absence of disproportionate edge-weighting in the input network (21, 25), but are found to be worth de-emphasizing further when optimizing overall learnability. Our findings are consistent with prior work showing that the difficulty of learning cross-cluster edges in modular networks is robust to the size and number of modules in the network (43).

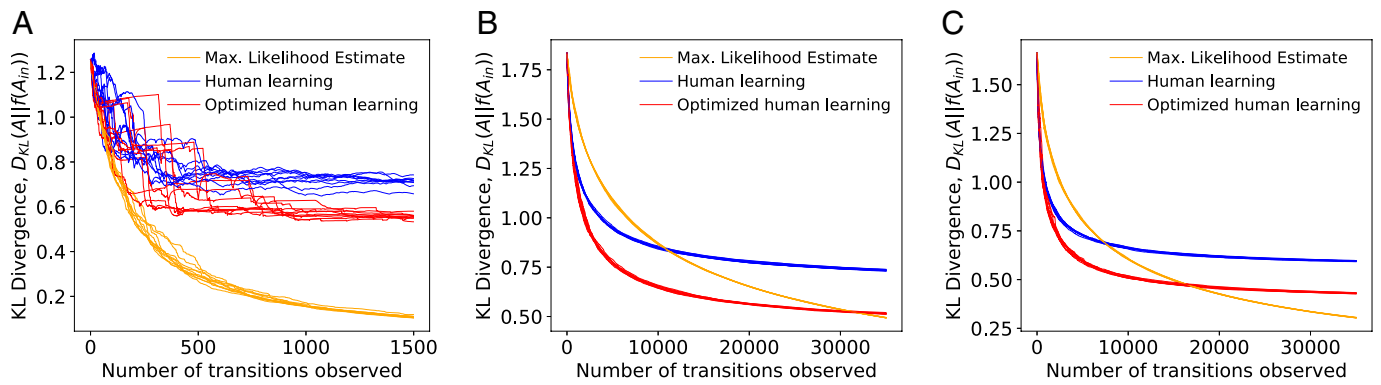


Fig. 5. Performance of network-learning strategies in the transient regime. Each panel shows the Kullback–Leibler (KL) divergence between some true network and the learned network as a function of the number of observed transitions. Three network-learning strategies are shown: maximum-likelihood estimation (optimal in the limit of infinite observations), standard human network learning (supported by ref. 21), and optimized human network learning (introduced in this paper). All plots report 10 simulations of each network-learning strategy, with human learning and optimized human learning simulations run at $\beta = 0.1$, close to the median learning accuracy reported in ref. 21. The three networks analyzed are the modular network with 15 nodes (A) (Fig. 1A), the semantic network extracted from the linear algebra textbook authored by Axler (B) (61), and the semantic network extracted from the linear algebra textbook authored by Edwards (C) (63).

Then, to probe whether our findings with the modular and lattice networks extended more generally to larger, more complex networks lacking a high degree of structural symmetry or uniform transition probabilities, we analyzed the optimization of learnability of networks generated from generative network models. We first began by studying the optimization of stochastic block networks. Importantly, we observed that for stochastic block networks with a high fraction of edges within communities, significant gains in network learnability can be achieved only by tuning a single parameter representing the weight of all cross-cluster edges. Specifically, we observed that stochastic block networks optimized for learnability de-emphasized cross-cluster edges. Next, motivated by the prevalence of heterogeneous degree distributions among real-world networks (37, 38, 44–46), we investigated the optimization of learning for degree-corrected stochastic block networks. Through applying a similar single-parameter optimization approach, we found that degree-corrected stochastic block networks share similar learning-optimization properties to standard stochastic block networks when tuning cross-cluster edge weights. However, the efficacy of these optimization strategies in improving network learnability was found to be slightly higher for degree-corrected stochastic block networks. This finding suggests that the learning of networks with hierarchically modular organization can be improved significantly more (using this cross-cluster edge-weight tuning) than can random modular networks. Taken together with prior work showing that hierarchically modular networks share similar information-theoretic properties with a large class of real-world networks (25) that random modular networks lack, these findings have implications for how features of real-world information networks ought to be weighted or designed.

Then, to understand optimal strategies for enhancing the learnability of real-world information networks, we analyzed how semantic networks extracted from college-level mathematics textbooks can be reweighted to maximize learnability. These networks exhibit core–periphery structure, indicating that they are composed of nodes that can roughly be divided into a densely connected core and a periphery that is loosely connected to nodes within the core (47–49). In addition, prior work has established that the periphery of these semantic networks possesses community structure (39). Importantly, unlike modular networks, we find that the semantic networks are not very optimizable near $\beta \approx 0$, but are significantly optimizable, even for moderately large values of β . One explanation for the difference in optimization near $\beta = 0$ between the modular networks and semantic networks

is that, for modular networks, optimal input networks approach disconnected graphs as $\beta \rightarrow 0$. Typically, as $\beta \rightarrow 0$, the learned representation from any input network approaches a uniform network with no particular structure (21, 25). However, a competing limit may occur when studying the optimization of learnability of modular networks, in that cross-cluster weights may also approach zero. In this limit, a network presentation approaches disconnected components, with each component representing a module of the original network. Thus, when β is taken close to zero, learning inaccuracies primarily strengthen edges within modules, causing only minor decreases in overall accuracy of the optimized network, as nodes within modules or clusters are already densely connected with each other in the original target network. In contrast, for networks without overall community structure, such as the semantic networks extracted from the mathematical texts, the optimal input graph always remains connected as $\beta \rightarrow 0$, and, thus, any learned representation will approach an all-to-all network with uniform transition probabilities in this limit.

To characterize the features of these semantic networks that become reinforced or de-emphasized when optimizing for learnability, we categorize the edges in each of the semantic networks by the core–periphery status of their endpoints, as well as by cross-cluster or within-cluster participation in periphery community structure. We find that edges contained solely among the periphery of these networks are emphasized the most, whereas core edges are de-emphasized, even at higher β . These findings suggest that the learnability of information networks can generally be enhanced by placing additional emphasis on relationships between less commonly occurring concepts and de-emphasizing highly central concepts. To assess this idea further, we investigated how the centrality measures of edge betweenness and edge degree were associated with the scaling of edge weights for these semantic networks. In particular, we found that both centrality measures are negatively correlated with optimal edge-weight scaling, confirming that, in these semantic networks, edges that are highly central generally should be de-emphasized to maximize learnability.

The optimization of networks for human learnability presents new angles for understanding real-world information networks. Recent work has demonstrated that science texts covering the same topic may vary greatly in their lexicon profiles and didactic approach (50). Understanding how edges in information networks associated with educational materials should be modulated may also be an insightful way of profiling the semantic content of

educational materials. Furthermore, while we have explored the optimization of semantic networks, other real-world information networks, such as phonological networks, may exhibit significantly different local and global topology (51). The erroneous clustering introduced in the human network-learning process would be especially detrimental for learning triangle-free substructures. Therefore, for some phonological networks, where star-like structures and leaf nodes are more common, one might expect that the network-optimization strategies employed in this work would de-emphasize edges to leaf nodes. Further investigation of how other kinds of information networks should be optimized may reveal new insights into network substructures that are optimal for learning.

Finally, we would like to draw attention to a possible connection between the ideas explored in this paper and the distinctions between human learning and machine learning. In particular, the simulations of network learning in the transient regime demonstrate that, although perfect, machine-like learning strategies outperform human learning after many observed transitions, human-learning strategies work surprisingly well in the low-data regime. This finding mirrors the observation that humans are capable of learning many tasks from few observations, in comparison to common machine-learning methods that often require large amounts of training examples to achieve decent performance. In view of this connection, an insightful way of explaining our results on transient-regime learning is that human learning strategies enforce helpful inductive biases in the learning process. Much like how regularization strategies in machine learning act as inductive biases and often improve performance in the low-data regime, the inductive biases of human learning greatly increase the data efficiency of the learning process. However, in both kinds of learning, when data are plentiful, inductive biases can hinder performance. While the present work primarily focuses on learning in the limit of infinite observations, future work should expand upon learning dynamics in the transient regime to uncover additional insights related to this connection.

Methodological Considerations. We note that our results on optimizing the learnability of generated networks are strictly a lower bound on improvements in human learnability of networks that can be afforded through targeted emphasis and de-emphasis of particular input network features. First, it is possible that relaxing the symmetry constraint enforced in the optimization process may lead to further improvements in network learnability. In addition, since we only consider optimization of learnability of generated networks via single-parameter tuning (cross-cluster edge weights or nonring edge weights), it is likely that more nuanced emphasis-modulation strategies may enhance learnability even further, beyond what has been demonstrated for both classes of stochastic block networks analyzed, as well as for Watts–Strogatz networks. Similarly, our results on semantic networks extracted from mathematics textbooks only represent lower bounds on improvements in learnability that can be achieved through targeted emphasis modulation. This is a consequence of the fact that each of the semantic networks analyzed had thousands of edge weights to be varied as parameters, and, thus, achieving globally optimal network representations via dual annealing was not always possible. In addition, the analyses presented in this work have yet to consider whether adding edges that are entirely nonexistent in a target network to the input network presented to a human learner may enhance learnability of the target network. Preliminary findings suggest that increases in learnability can indeed occur when nonexistent edges are added to an input network representation (*SI Appendix, Fig. S5*). Future work could fruitfully consider how

relaxing edge-existence constraints may affect the efficacy of the estimated optimal network representations.

Conclusion

Recent advances in the study of human information processing have shed light on the ways that humans learn information networks. Rather than mapping the structure of networks exactly, inaccuracies in human learning often result in erroneous or biased internal representations. To overcome inaccuracies in human learning, we investigate how networks presented to human learners can be designed to counteract and minimize inaccuracies in learning. Across a range of synthetic networks, we find that reinforcing edges within clusters and de-emphasizing edges between clusters improves network learnability. In addition, we analyze how real-world semantic networks can be optimized for learnability, thereby uncovering the fact that relationships between periphery concepts ought to be reinforced. Together, our findings demonstrate that the learnability of network representations can be significantly enhanced through intentionally modulating the emphasis of specific network features.

Materials and Methods

Converting between Weighted Graphs and Transition Matrices. To convert between graphs and transition matrices, we use the following two relations:

Given a weighted graph G , the normalized transition matrix corresponding to random walks is given by $A_{ij} = \frac{G_{ij}}{\sum_{k=1}^N G_{ik}}$.

Conversely, the adjacency matrix for a weighted, undirected graph corresponding to a reversible transition matrix A is determined up to a multiplicative constant by $G_{ij} = \pi_i A_{ij}$, where π is the stationary distribution of A .

Optimization Methods. To assess how much a learned network differs from the target network, we use the Kullback–Leibler divergence. The Kullback–Leibler divergence between normalized transition networks A and B is defined as

$$D_{KL}(A||B) = - \sum_i \pi_i \sum_j A_{ij} \log \left(\frac{B_{ij}}{A_{ij}} \right), \quad [2]$$

where π is the stationary distribution of A , and only terms with nonzero transition probabilities are summed over.

For some target transition structure A , the optimal input structure $A_{in} = A^*$ was determined by using the dual-annealing optimization method in *scipy*, with $D_{KL}(A||f(A_{in}))$ as the cost function. During the optimization process, the edge weights of A_{in} corresponding to edges that exist in A were varied as free parameters bounded between zero and one. The matrix A_{in} was then normalized prior to every cost-function evaluation. For the modular and lattice graphs, the number of free parameters varied during optimization was reduced based on the networks' respective symmetries. Network symmetries were computed by using the *iGraph* package in Python (52).

Network Generation. We generated stochastic block networks as follows: Starting with N nodes, we first assigned each node to a community labeled 1 through 5 uniformly at random. Then, $\lfloor |E| \rho \rfloor$ edges were added between nodes in the same module, where $|E|$ is the total number of edges in the network. Specifically, each addition of a within-cluster edge was performed by selecting a community uniformly at random, selecting two nonadjacent nodes within the community uniformly at random, and then adding an edge between the selected nodes. Finally, $|E| - \lfloor |E| \rho \rfloor$ edges were then added between nodes in different communities. This process was performed by selecting two different communities uniformly at random, then selecting one node in each community uniformly at random, and connecting the nodes if they were nonadjacent. If they were adjacent, the step was repeated.

For degree-corrected stochastic block networks, a similar network-generation method was used, with modifications adapted from the procedure described in ref. 25. Specifically, each node was assigned an index from i from 1 to 200 and was then assigned a weight i^{-1} . Then, the probability of a given node being chosen for the addition of an edge at any step was proportional to the weight of the node.

Analysis of Semantic Networks. Given a weighted network $G = (V, E)$ with edge weights w_{ij} , we identified core-periphery structure by finding a partition of the vertex set V into disjoint sets C and P so that the following core-ness quality function is maximized:

$$Q_C = \frac{1}{v_C} \left(\sum_{ij \in C} (w_{ij} - \gamma_C \bar{w}) - \sum_{ij \in P} (w_{ij} - \gamma_C \bar{w}) \right), \quad [3]$$

where v_C is a normalization constant, \bar{w} is the average over all edge weights, and γ_C is a resolution parameter that we set to one.

Given a partition $V = C \cup P$ into core and periphery nodes, we evaluate the community structure of the periphery of G by maximizing the following modularity quality function on the subgraph of G induced on P :

$$Q_C = \frac{1}{v_M} \left(\sum_{ij \in P} (w_{ij} - \gamma_M \frac{s_i s_j}{v_M}) \delta_{ij} \right), \quad [4]$$

where v_M is a normalization constant, $s_i = \sum_{j \in P} w_{ij}$ is the sum over the weights of all edges incident on vertex i , and γ_M is a resolution parameter that we set to one.

The edge-betweenness centrality for some edge e is defined as

$$C_B(e) = \frac{2}{(n-1)(n-2)} \sum_{s,t \in V} \frac{\sigma(s,t|e)}{\sigma(s,t)}, \quad [5]$$

where $\sigma(s,t)$ is the number of weighted shortest paths between nodes s and t , and $\sigma(s,t|e)$ is the number of such shortest paths that go through the edge e .

For an edge $e = \{i,j\}$, we define the weighted edge degree centrality as

$$C_D(e) = \frac{s_i + s_j}{2 \sum_{p < q} w_{pq}}, \quad [6]$$

where $s_i = \sum_k w_{ik}$ is the weighted degree of node i in the network.

Simulating Transient Network Learning. While observing random walks drawn from some transition network, the maximum-likelihood estimate for the transition network can be described by

$$A_{ij}^{\text{MLE}}(t) = \frac{n_{ij}(t)}{\sum_k n_{ik}(t)}, \quad [7]$$

where n_{ij} is the number of observed transitions from node i to node j by time step t . As described in ref. 21, the human learned representation takes a similar form:

$$\tilde{A}_{ij}(t) = \frac{\tilde{n}_{ij}(t)}{\sum_k \tilde{n}_{ik}(t)}, \quad [8]$$

- J. R. Saffran, R. N. Aslin, E. L. Newport, Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
- J. Fiser, R. N. Aslin, Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* **28**, 458–467 (2002).
- D. Perani et al., Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16056–16061 (2011).
- A. E. Sizemore, E. A. Karuza, C. Giusti, D. S. Bassett, Knowledge gaps in the early growth of semantic feature networks. *Nat. Hum. Behav.* **2**, 682–692 (2018).
- E. A. Karuza, S. L. Thompson-Schill, D. S. Bassett, Local patterns to global architectures: Influences of network topology on human learning. *Trends Cogn. Sci.* **20**, 629–640 (2016).
- T. Engelthaler, T. T. Hills, Feature biases in early word learning: Network distinctiveness predicts age of acquisition. *Cogn. Sci. (Hauppauge)* **41** (suppl. 1), 120–140 (2017).
- S. H. Solomon, J. D. Medaglia, S. L. Thompson-Schill, Implementing a concept network model. *Behav. Res. Methods* **51**, 1717–1736 (2019).
- M. Peer, I. K. Brunec, N. S. Newcombe, R. A. Epstein, Structuring knowledge with cognitive maps and cognitive graphs. *Trends Cogn. Sci.* **25**, 37–54 (2021).

where \tilde{n}_{ij} is the revised (and erroneous) count of transitions from node i to node j . In particular, it is described by

$$\tilde{n}_{ij}(t+1) = \tilde{n}_{ij}(t) + B_t(i)[j = x_{t+1}], \quad [9]$$

where x_{t+1} is the node observed at time step $t+1$, $[\cdot]$ is the Iverson bracket, and $B_t(i)$ encodes an internal belief about which node was observed at time t , described by

$$B_t(i) = \frac{1}{Z} \sum_{\Delta t=0}^{t-1} e^{-\beta \Delta t} [i = x_{t-\Delta t}], \quad [10]$$

where Z represents a normalizing constant.

Citation Diversity Statement. Recent work in several fields of science has identified a bias in citation practices, such that papers from women and other minorities are undercited relative to the number of such papers in the field (53–58). Here, we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, and other factors. We obtained predicted gender of the first and last author of each reference by using databases that store the probability of a name being carried by a woman (53, 59). By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 21.05% woman(first)/woman(last), 7.89% man/woman, 14.17% woman/man, and 56.88% man/man. This method is limited in that 1) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity; and 2) it cannot account for intersex, nonbinary, or transgender people. We look forward to future work that could help us to better understand how to support equitable practices in science.

Data, Materials, and Software Availability. Previously published data were used for this work from GitHub (<https://github.com/nhchristianson/Math-text-semantic-networks>). All code and data used in analyses have been made publicly available at <https://github.com/wqjan0/OptimizedGraphLearning> (60).

ACKNOWLEDGMENTS. We thank Christopher Kroninger for feedback on earlier versions of this manuscript. We also thank Pixel Xia for useful discussions on Sierpiński graph construction. This work was supported by the Army Research Office (DCIST-W911NF-17-2-0181) and the National Institute of Mental Health (1-R21-MH-124121-01). D.S.B. acknowledges additional support from the John D. and Catherine T. MacArthur Foundation, the Institute for Scientific Interchange Foundation, NSF CAREER Award PHY-1554488, and the Center for Curiosity. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Author affiliations: ^aDepartment of Physics and Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA 19104; ^bInitiative for the Theoretical Sciences, Graduate Center, City University of New York, New York, NY 10016; ^cJoseph Henry Laboratories of Physics, Princeton University, Princeton, NJ 08544; ^dLewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544; ^eDepartment of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104; ^fDepartment of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125; ^gDepartment of Electrical and Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA 19104; ^hDepartment of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; ⁱDepartment of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; and ^jSanta Fe Institute, Santa Fe, NM 87501

- E. J. Timothy et al., What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018).
- R. V. Sole, B. Corominas-Murtra, S. Valverde, L. Steels, Language networks: Their structure, function, and evolution. *Complexity* **15**, 20–26 (2010).
- X. F. Liu, K. T. Chi, M. Small, Complex network structure of musical compositions: Algorithmic generation of appealing music. *Phys. A* **389**, 126–132 (2010).
- G. A. Miller, C. Fellbaum, Semantic networks of English. *Cognition* **41**, 197–229 (1991).
- S. C. Herring et al., "Language networks on livejournal" in *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)* (IEEE, Piscataway, NJ, 2007), p. 79.
- I. Peretz et al., Music lexical networks: The cortical organization of music recognition. *Ann. N.Y. Acad. Sci.* **1169**, 256–265 (2009).
- A. C. Schapiro, T. T. Rogers, N. I. Cordova, N. B. Turk-Browne, M. M. Botvinick, Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).
- A. C. Schapiro, N. B. Turk-Browne, K. A. Norman, M. M. Botvinick, Statistical learning of temporal community structure in the hippocampus. *Hippocampus* **26**, 3–8 (2016).
- S. H. Tompson, A. E. Kahn, E. B. Falk, J. M. Vettel, D. S. Bassett, Functional brain network architecture supporting the learning of social networks in humans. *NeuroImage* **210**, 116498 (2020).

18. J. Stiso *et al.*, Neurophysiological evidence for temporally discounted cognitive map formation during sequence learning. *bioRxiv* [Preprint] (2021). <https://www.biorxiv.org/content/10.1101/2021.09.07.459302v1>. Accessed 12 August 2022.
19. J. W. Rueckemann, M. Sosa, L. M. Giocomo, E. A. Buffalo, The grid code for ordered experience. *Nat. Rev. Neurosci.* **22**, 637–649 (2021).
20. C. W. Lynn, D. S. Bassett, How humans learn and represent networks. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 29407–29415 (2020).
21. C. W. Lynn, A. E. Kahn, N. Nyema, D. S. Bassett, Abstract representations of events arise from mental errors in learning and memory. *Nat. Commun.* **11**, 2313 (2020).
22. I. Momennejad *et al.*, The successor representation in human reinforcement learning. *Nat. Hum. Behav.* **1**, 680–692 (2017).
23. K. L. Stachenfeld, M. M. Botvinick, S. J. Gershman, The hippocampus as a predictive map. *Nat. Neurosci.* **20**, 1643–1653 (2017).
24. P. Dayan, Improving generalisation for temporal difference learning: The successor representation. *Neural Comput.* **5**, 613–624 (1993).
25. C. W. Lynn, L. Papadopoulos, A. E. Kahn, D. S. Bassett, Human information processing in complex networks. *Nat. Phys.* **16**, 965–973 (2020).
26. R. Haven Wiley, *Noise Matters: The Evolution of Communication* (Harvard University Press, Cambridge, MA, 2015).
27. L. Real, *Behavioral Mechanisms in Evolutionary Ecology* (University of Chicago Press, Chicago, 1994).
28. E. A. Karuza, A. E. Kahn, S. L. Thompson-Schill, D. S. Bassett, Process reveals structure: How a network is traversed mediates expectations about its architecture. *Sci. Rep.* **7**, 12733 (2017).
29. A. E. Kahn, E. A. Karuza, J. M. Vettel, D. S. Bassett, Network constraints on learnability of probabilistic motor sequences. *Nat. Hum. Behav.* **2**, 936–947 (2018).
30. S. H. Tompson, A. E. Kahn, E. B. Falk, J. M. Vettel, D. S. Bassett, Individual differences in learning social and nonsocial network structures. *J. Exp. Psychol. Learn. Mem. Cogn.* **45**, 253–271 (2019).
31. V. Bakir, E. Herring, D. Miller, P. Robinson, Organized persuasive communication: A new conceptual framework for research on public relations, propaganda and promotional culture. *Crit. Sociol.* **45**, 311–328 (2019).
32. A. W. Bloedel, I. Segal, Persuasion with rational inattention. SSRN [Preprint] (2018). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3164033. Accessed 12 August 2022.
33. R. Boleslavsky, K. Kim, Bayesian persuasion and moral hazard. SSRN [Preprint] (2017). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913669. Accessed 12 August 2022.
34. C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **52**, 479–487 (1988).
35. C. Tsallis, D. A. Stariolo, Generalized simulated annealing. *Physica A* **233**, 395–406 (1996).
36. Y. Xiang, D. Y. Sun, W. Fan, X. G. Gong, Generalized simulated annealing algorithm and its application to the Thomson model. *Phys. Lett. A* **233**, 216–220 (1997).
37. K. Lerman, X. Yan, X.-Z. Wu, The “majority illusion” in social networks. *PLoS One* **11**, e0147617 (2016).
38. A.-L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
39. C. H. Christianson, A. Sizemore Blevins, D. S. Bassett, Architecture and evolution of semantic networks in mathematics texts. *Proc. Math. Phys. Eng. Sci.* **476**, 20190741 (2020).
40. N. Christianson, Semantic networks in mathematics texts [Jupyter Notebook]. GitHub. <https://github.com/nhchristianson/Math-text-semantic-networks>. Deposited 6 August 2020.
41. D. C. McNamee, K. L. Stachenfeld, M. M. Botvinick, S. J. Gershman, Flexible modulation of sequence generation in the entorhinal-hippocampal system. *Nat. Neurosci.* **24**, 851–862 (2021).
42. C. R. James *et al.*, The Tolman-Eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell* **183**, 1249–1263.e23 (2020).
43. K. A. Karuza, A. E. Kahn, D. S. Bassett, Human sensitivity to community structure is robust to topological variation. *Complexity* **2019**, 8379321 (2019).
44. M. Nekovee, Y. Moreno, G. Bianconi, M. Marsili, Theory of rumour spreading in complex social networks. *Phys. A* **374**, 457–470 (2007).
45. R. Albert, Scale-free networks in cell biology. *J. Cell Sci.* **118**, 4947–4957 (2005).
46. B. Tadić, M. Mitrović, Jamming and correlation patterns in traffic of information on sparse modular networks. *Eur. Phys. J. B* **71**, 631–640 (2009).
47. P. Csermely, A. London, L.-Y. Wu, B. Uzzi, Structure and dynamics of core/periphery networks. *J. Complex Netw.* **1**, 93–123 (2013).
48. T. Verma, F. Ruffmann, N. A. M. Araújo, J. Nagler, H. J. Herrmann, Emergence of core-peripheries in networks. *Nat. Commun.* **7**, 10441 (2016).
49. M. P. Rombach, M. A. Porter, J. H. Fowler, P. J. Mucha, Core-periphery structure in networks. *SIAM J. Appl. Math.* **74**, 167–190 (2014).
50. I. T. Koponen, M. Nousiainen, *Lexical Networks and Lexicon Profiles in Didactical Texts for Science Education* (Studies in Computational Intelligence, Springer International Publishing, Cham, Switzerland, 2020), vol. 882, pp. 15–27.
51. K. D. Neergaard, J. Luo, C.-R. Huang, Phonological network fluency identifies phonological restructuring through mental search. *Sci. Rep.* **9**, 15984 (2019).
52. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal Complex Syst.* **1695**, 1–9 (2006).
53. J. D. Workin *et al.*, The extent and drivers of gender imbalance in neuroscience reference lists. *Nat. Neurosci.* **23**, 918–926 (2020).
54. D. Maliniak, R. Powers, B. F. Walter, The gender citation gap in international relations. *Int. Organ.* **67**, 889–922 (2013).
55. N. Caplar, S. Tacchella, S. Birrer, Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nat. Astron.* **1**, 0141 (2017).
56. P. Chakravarty, R. Kuo, V. Grubbs, C. McIlwain, #communicationswhite. *J. Commun.* **68**, 254–266 (2018).
57. Y. Thiem, K. F. Sealey, A. E. Ferrer, A. M. Trott, R. Kennison, “Just ideas? The status and future of publication ethics in philosophy: A white paper” (Tech. Rep., Publication Ethics, Fairfield University, Fairfield, CT, 2018).
58. M. L. Dion, J. L. Sumner, S. M. Mitchell, Gendered citation patterns across political science and social science methodology fields. *Polit. Anal.* **26**, 312–327 (2018).
59. D. Zhou *et al.*, Gender diversity statement and code notebook v1.0. Zenodo (2020). <https://doi.org/10.5281/zenodo.3672110>. Accessed 12 August 2022.
60. W. Qian, Optimized Graph Learning. GitHub. <https://github.com/wqian0/OptimizedGraphLearning>
61. S. Axler, *Linear Algebra Done Right*, Third Edition (Springer International Publishing, 2015).
62. O. Bretscher, *Linear Algebra with Applications*, Fifth Edition (Pearson Education, Inc., 2013).
63. H. M. Edwards, *Linear Algebra* (Springer Science+Business Media, 1995).
64. W. H. Greub, *Linear Algebra*, Third Edition (Springer Science+Business Media, 1967).
65. J. Hefferson, *Linear Algebra*, Third Edition (2017). <http://joshua.smcvt.edu/linearalgebra>. Accessed 12 August 2022.
66. S. Lang, *Introduction to Linear Algebra*, Second Edition (Springer Science+Business Media, 1986).
67. P. Petersen, *Linear Algebra* (Springer Science+Business Media, 2012).
68. L. Robbiano, *Linear Algebra for everyone* (Springer-Verlag Italia, 2011).
69. G. Strang, *Linear Algebra and Its Applications*, Fourth Edition (Thomson Learning, 2006).
70. S. Treil, *Linear algebra done wrong* (2017). <https://www.math.brown.edu/~treil/papers/LADW/LADW.html>. Accessed 12 August 2022.