Cognitive network enrichment not degradation explains the aging mental lexicon and links fluid and crystallized intelligence

Thomas T. Hills[1]

[1] University of Warwick

Author Note

This work was supported by the Alan Turing Institute and Royal Society Wolfson Research Merit Award WM160074. The R code to reproduce the environments, the learning of cognitive representations, the behavior, and all analyses and figures in the manuscript and Supplementary Material are available at https://github.com/thomasthills/enrichment. The concept for this manuscript was initially proposed as a network example in Hills (2025) textbook. This research was not preregistered. Thanks to members of the Warwick Psychology Department and Warwick Business School for helpful comments.

Correspondence concerning this article should be addressed to Thomas T. Hills, Gibbet Hill Road, Coventry, CV4 7AL, UK. E-mail: t.t.hills@warwick.ac.uk

Abstract

Cognition is a complex system of interacting components. Late-life cognitive decline is often explained as a degradation of the interconnectivity among these components. Evidence from the aging mental lexicon corroborates this interpretation as older adults produce higher entropy responses in free association tasks, appear to have sparser free association networks, and judge objects to be less similar to one another than younger adults. Here, I demonstrate that all of these effects are produced by a model of cognitive network enrichment, which treats aging as an extension of lifelong learning. By increasing interconnectivity, learning increases competition for activation among potential targets, increasing entropy and reducing targeted activation. The impact of network enrichment is demonstrated using a general prediction error model (Rescorla-Wagner) which learns and enriches a cognitive network representation following lifelong experience with a network of associations in the environment. Sampling from the learned representation to produce behavior reproduces the above effects. A qualitative model comparison shows that various models of degradation fail to capture the above results for entropy and similarity. Both enriched and degraded representations can produce sparsening free association networks, depending on the specific methodological details of data collection. This underscores the general problem of inferring representation from behavior without considering process. Further, extending cognitive network enrichment more broadly provides a lifelong developmental pathway for over-attention to irrelevant stimuli and cognitive slowing—increasing interference, taxing resource limitations, and reducing targeted activation—offering a common cause for rising crystallized intelligence and declining fluid intelligence.

Keywords: cognitive aging; Rescorla Wagner; network science; free associations; fluid intelligence; crystallized intelligence

Cognitive network enrichment not degradation explains the aging mental lexicon and links fluid and crystallized intelligence

Cognitive aging is characterized by two distinct and well-documented patterns. As individuals age, measures of working memory, processing speed, and long-term memory show performance decrements from approximately age 20, while at the same time, measures of vocabulary and other kinds of general knowledge continue to increase (Brysbaert, Stevens, Mandera, & Keuleers, 2016; D. M. Burke & Shafto, 2004; Hartshorne & Germine, 2015; Keuleers, Stevens, Mandera, & Brysbaert, 2015; Park & Reuter-Lorenz, 2009; Salthouse, 2004, 2019; Shelton, Elliott, Matthews, Hill, & Gouvier, 2010; Troyer, Moscovitch, & Winocur, 1997; Verhaeghen, 2003). Though much debate surrounds the general classification of these abilities, their divergence corresponds to a classic division of intelligence between the ability to solve novel problems in a fast and accurate way, called *fluid intelligence*, and the quantity of one's prior knowledge, called *crystallized intelligence* (Cattell, 1987; Horn, 1989). This division also characteristically distinguishes the old from the young.

Longitudinal studies find that healthy declines in fluid intelligence are predominantly correlated with rising crystallized intelligence (Tucker-Drob et al., 2022). Might this divergence share a common cause? As described below, many researchers have postulated such a relationship and the weight of the evidence is growing. However, a general model demonstrating this relationship as a consequence of life-experience remains to be formally proposed. The model of cognitive network enrichment that I propose integrates well-established mechanisms of learning and memory across a lifetime of experience. This integration involves modeling the environment, memory encoding as a result of experience with that environment, and the behavior that arises from that encoding. The results of this developmental process demonstrate a clear interaction: enriching one's cognitive representation increases competition between associates during retrieval and produces a relative decline in activation between any two concepts chosen at random.

These results are robust to a wide variety of environments and learning processes. To prepare the reader, the model I propose is not a model of working memory—models demonstrating the influence of interference and competition in working memory already exist (e.g., Oberauer, Farrell, Jarrold, & Lewandowsky, 2016; Oberauer & Lin, 2024) and older adults suffer more interference from prior knowledge (e.g., Amer, Wynn, & Hasher, 2022). Rather, cognitive network enrichment demonstrates how an enriched cognitive representation enhances competition for access to working memory and dilutes retrieval whenever prior knowledge comes into play. These results offer an enrichment-based explanation for aging effects commonly associated with memory degradation, which, as I show below, degradation does not produce. Before demonstrating how these effects arise, I first highlight the primary theoretical accounts of aging this work addresses and then describe the key behavioral effects I hope to explain.

Degradation or enrichment

Cognitive aging is a rich process with evidence supporting a wide range of phenomenology and theoretical explanations (Grady, Springer, Hongwanishkul, McIntosh, & Winocur, 2006; Koen & Rugg, 2019; Mata, Schooler, & Rieskamp, 2007; Reuter-Lorenz & Cappell, 2008; Spreng & Turner, 2021; Troyer et al., 1997). Much of this work is interpreted as supporting either degradation or enrichment.

One prominent degradation-based explanation is the common cause theory of age-related cognitive decline, which argues that biological aging in the brain is the source of processing speed deficits (Baltes & Lindenberger, 1997; Deary et al., 2009). The supposition is that aging is a general process of degradation, in which factors like oxidative stress and telomere shortening damage the physiological architecture underpinning cognitive performance. Salthouse (1992) describes some potential mechanisms as follows:

"a slower speed of transmission along single (e.g., loss of myelination) or multiple (e.g., loss of functional cells dictating circuitous linkages) pathways, or. . . delayed propagation at the connections between neural units (e.g., impairment in functioning of neurotransmitters, reduced synchronization of activation patterns)" (p. 116). Consistent with this, percent volume of grey and white-matter declines in late life (Ge et al., 2002; Giorgio et al., 2010; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003) as does cortical thickness (Lemaitre et al., 2012), even if cell death is not characteristic of healthy aging (S. N. Burke & Barnes, 2006).

Evidence of apparent degradation extends to studies of brain-wide integration. For example, measures of the neural connectome—the wiring diagram of the brain—indicate declining local community structure and a reduction in functional segregation (Cao et al., 2014; Damoiseaux, 2017; Ferreira et al., 2016; Riedel, Heuvel, & Markett, 2022). This dedifferentiation is marked by "reduced within-network and increased between-network functional connectivity" (Deery, Di Paolo, Moran, Egan, & Jamadar, 2023). Evidence that this dedifferentiation is caused by degradation is however mixed. Across studies, brain atrophy frequently explains limited variance in functional desegregation and the two are often statistically uncorrelated (Ferreira et al., 2016; Geerligs, Renken, Saliasi, Maurits, & Lorist, 2015). The limited nature of this relationship between degradation and decline is found elsewhere as well, even in more direct studies. Posthumous evidence of Alzheimer's and other neurodegenerative and cerebrovascular diseases account for around 40% of declines in fluid intelligence (e.g., the Mini-Mental State Examination), leaving substantial variance unexplained even in unhealthy individuals (Boyle et al., 2021).

An alternative explanation for age-related cognitive decline proposes a causal interdependence between crystallized and fluid intelligence. More specifically, learning is proposed to increase crystallized intelligence but impair fluid intelligence. Buchler and Reder (2007) showed using simulations that if one assumes that the number of relations between concepts increases as a result of learning over the lifespan this would directly lead to more diffuse activation between those concepts. As knowledge increases, activation is more broadly distributed. Similarly, Ramscar, Hendrix, Shaoul, Milin, and Baayen (2014) demonstrated how prior learning explains age-related declines in paired-associate learning. They based their work on Desrosiers and Ivison (1986), which found that older adults perform most poorly on word pairs that are least consistent with prior experience. The difficulty of learning unrelated word pairs is entirely predictable from the co-occurrence frequency of those pairs in prior experience (see also Ramscar, Hendrix, Love, & Baayen, 2013). Training a Rescorla-Wagner model on typical patterns of word co-occurrences leads unrelated word pairs to become negatively associated over time, impairing their future learning. This negative association, formally called conditioned inhibition, is a predictable but often understated property of Pavlovian conditioning (Rescorla, 1988).

Still more recent work has argued for a much broader influence of age-related mental "clutter," which may arise from representational changes across the lifespan as well as changes in cognitive control at the time of encoding or retrieval (Amer et al., 2022; Weeks, Grady, Hasher, & Buchsbaum, 2020). According to this account, over-attending to a diversity of stimuli alongside an inability to filter out past experience can lead older adults to attend to too much information (e.g., Lustig, May, & Hasher, 2001). Too much information creates processing difficulties, further exacerbated when that information is irrelevant.

One source of this informational clutter is life experience. Li and Siew (2022) found that older individuals spent more time processing words that had changed their meaning during their lifetime, a delay not found in younger individuals who were too young to have experienced the change. Similarly, Hoffman (2019) used a verbal comprehension test involving synonym identification and found that the performance of older adults was most poor when the questions contained relatively strong semantic distractors, a relationship not found for younger individuals. Numerous studies like these show that prior knowledge influences older adults more strongly than younger adults (see Spreng & Turner, 2019), and this may be because there is more of it. Though one could argue that some of these results are due to putative age-related deficits in cognitive control, the work from Li and Siew (2022) does not support that argument—older adults are similar to younger adults on words that have not changed over their lifetime.

The different accounts described above can be broadly categorized as focusing on degradation or enrichment, and the evidence for both is compelling. To what extent we subscribe to one explanation over another should largely depend on the plausibility of their mechanisms and the sufficiency of what they explain. To my knowledge, degradation accounts have not provided formal computational mechanisms for how degradation might lead to the observed age-related changes in healthy individuals. Presumably such mechanisms could be developed and would offer useful predictions. For example, Borge-Holthoefer, Moreno, and Arenas (2011) developed a model of degradation for hyper-priming in Alzheimer's patients, but similar models for healthy aging are still needed (see also Stella, 2020). A key challenge for degradation models is that they would need to explain how degradation alters fluid intelligence but not crystallized intelligence. I offer two such models that degrade associations between concepts without altering the persistence of those concepts. Neither of them capture the behavioral effects I describe below.

On the other hand, existing enrichment models have either assumed more connectivity (e.g., Buchler & Reder, 2007) or evaluated how differential experience to word pairs impairs future learning (e.g., Ramscar et al., 2014). As noted by Wulff, De Deyne, Jones, and Mata (2019), what is lacking is a full model of representational development and behavior across the adult lifespan. The cognitive network enrichment account I propose here is one such model based solely on extending the natural learning process. It represents a computational prediction for what we should expect if late-life were a continuation of early life. In addition, this enrichment account also helps identify what remains to be explained by degradation and provides a platform for testing degradation's impacts.

To evaluate these models, I focus on two empirical observations made over the last decade. These tap into different aspects fundamental to cognitive retrieval: entropy in free associations and judgments of similarity. These address how the mind accesses knowledge in divergent and convergent ways and offer tasks for directly evaluating how structure and process influence behavior. In explaining how enrichment alone alters the mental lexicon and gives rise to the empirical results associated with the above phenomena, we can see how lifelong experience influences divergent and convergent features of cognitive processing, two fundamental components of fluid intelligence.

The aging mental lexicon: rising entropy and falling similarity

The fate of the enrichment and degradation accounts rests entirely on their ability to accurately describe age-related change in the structure of cognitive representations. Because we do not have direct access to these representations, we must infer them from behavior. Numerous studies have attempted to do this. Dubossarsky, De Deyne, and Hills (2017) examined data from approximately 8000 people, ranging in age from roughly 10 to 70, who provided three free associates to each of 420 cue words. Data were aggregated within age-groups to evaluate changes in associations across the lifespan. From approximately age 30, the entropy of associations increased—associates become less predictable with increasing age. Though this rising entropy could be a consequence of aggregation across a more diverse older population, Jeong and Hills (2024) showed that this was a property of individuals. Older individuals produce successive associates in response to a single cue that are less similar to one another than do younger individuals.

Dubossarsky et al. (2017) also produced networks among the 420 cue words, with weighted edges connecting cues more strongly when they produced more similar associations. Older networks were found to be more sparsely connected. They had a lower average degree (number of associations per word) higher average shortest path length (greater path distance between associates), and lower average clustering coefficient (the proportion of a concept's neighbors that are themselves connected). Both Zortea, Menegola, Villavicencio, and Salles (2014) and Wulff, De Deyne, Aeschbach, and Mata (2022) identified similar patterns of declining connectivity with varying numbers of participants and cues. Jeong and Hills (2024) further corroborated these results for three different languages (English, Dutch, and Spanish).

These changes in the aging lexicon are also consistent with the way older and younger adults search memory. Using a semantic fluency task (e.g., "name all the animals you can"), Wulff, Hills, and Mata (2022) constructed lexical networks by connecting animal names that appeared nearby one another in the fluency lists of older and younger adults. Older adults' lexical networks were less well-connected, similar to the patterns for free associations described above. This poor connectivity was also observed by fitting a memory search model to semantic fluency data from younger and older adults. The model found that, compared with younger adults, older adults produced strings of animal names that were less well predicted by semantic similarity between names (Hills, Mata, Wilke, & Samanez-Larkin, 2013). Finally, Cosgrove, Kenett, Beaty, and Diaz (2021) used percolation analysis to investigate the resilience of networks built from semantic fluency data from older and younger individuals. By artificially removing connections between words, they found that older adult networks were less resilient to decay than younger adult networks.

Similarity judgments show a similar divergence: Older adults tend to rate things as less similar to one another than younger adults. Wulff, Hills, et al. (2022) demonstrated this by asking younger and older adults to judge the similarity of 63 common animals, making approximately 2000 paired comparisons over a period of several weeks using a tablet participants took to their homes. Using a similar task, Cosgrove, Beaty, Diaz, and Kenett (2023) found a corresponding decline in similarity judgments with age. They also found that older individual networks produced from similarity judgments were less well connected, with lower clustering coefficients and lower efficiency (a measure of inverse path length). Foreshadowing the results of cognitive network enrichment, Cosgrove et al. (2023)

also found that larger vocabularies were associated with lower clustering coefficients and global efficiency, suggesting that cognitive enrichment may play a role in poorer connectivity.

In sum, with age adults become less predictable in free associations tasks (higher entropy), produce judgments of lower similarity in pairwise comparisons, and appear to have a sparsening free association network (lower degree, lower clustering coefficient, and higher average shortest path length). These results are intuitively consistent with a cognitive representation that suffers from degradation. One can easily imagine that a sparseness in output reflects a sparseness in the underlying representation, which is caused by degradation in the underlying biological architecture. However, as I show below, this intuition is wrong. It is wrong because it conflates the outputs of a retrieval process with the structure of the underlying representation. When we separate the behavior from the representation and the processes that give rise to them we find that extending standard learning and retrieval models across the lifespan predicts all of the above effects.

Cognitive Network Enrichment

Cognitive network enrichment takes its inspiration from complex systems, in which behavior is the emergent outcome of self-organizing and distributed processes acting across interconnected components (Simon, 1977; Strogatz, 2001). In cognitive science, as in other complex systems, this is often manifested in how network structure shapes processing and subsequent behavior (Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013; Hills, 2025; Siew, Wulff, Beckage, & Kenett, 2019). The networks commonly used to characterize the mental lexicon are themselves constructed from behavioral outputs (e.g., free associations tasks). They are therefore underpinned by both structure and process. As a consequence, to understand how the mental lexicon ages, it is essential to carefully consider the processes by which the lexicon is formed and accessed.

The cognitive network enrichment account envisions lifelong behavior as the ongoing outcome of learning from the environment to develop a cognitive representation, and then applying processes to that representation to generate behavior. This enrichment account does not assume that the representation is a perfect copy of the environment, nor that behavior is a perfect copy of the representation. Each requires processes to transform one into the other. This approach follows calls from cognitive research community to model not only the representation but the processes that access that representation to generate behavior (Castro & Siew, 2020; Estes, 1975; Hills & Kenett, 2022; Johns, Jamieson, & Jones, 2023; Jones, Hills, & Todd, 2015)

Formally, the cognitive network enrichment account involves modeling three separate components:

1. *Environment*: The environment presents the set of all possible associations that could be learned.
2. *Representation*: The representation is the outcome of learning associations from the environment, with learning continuing to enrich the representation across the lifespan.
3. *Behavior*: Behavior is guided by retrieving information from the cognitive representation appropriate to the environmental context. This generates free-associations, memory search, similarity judgments, and so on.

To make this relationship clear, these stages are presented in Figure 1 and each are explained in detail below. Though there is nothing controversial about this framing, it is visualized here because, as Hills and Kenett (2022) and Jones et al. (2015) have noted, it is easy to conflate the behavior with the representation or the representation with the environment. The failure to separate these components limits our intuition for understanding cognition as an outcome of both representation and process. Free associations should not be confused with a copy of the underlying representation. Similarity judgments should not be confused with indicating the distance between objects in memory. In both cases, they are the output of applying a process to a representation. Even when the processes remain unchanged, changes in the structure can influence the output in misleading ways, and vice versa (Hills, 2025; Hills & Kenett, 2024). To provide a simple analogy: That needles are harder to find in bigger haystacks does not mean that there are fewer needles. Nor does it mean that we are less capable of searching for them.

Figure 1. The cognitive network enrichment account models the process of translating experience with the environment into behavior. Arrows represent processes that translate one domain into another. Learning (e.g., Rescorla-Wagner) translates experience with the environment into a cognitive representation. Additional cognitive processes (e.g., spreading activation) translate the representation into behavior.

Environment

The environment is represented as a network of concepts (e.g., nodes) connected by weighted edges (links). Edges represent associations that can be learned and their weight indicates the intensity of the association. This follows the basic idea that things in the world are related to one another to varying degrees and in experiencing the world we learn those relationships. This applies to language, such as co-locations of words, but it applies equally well to any set of entities in the world about which things can be learned. I refer to these things as concepts to reflect their generality.

The enrichment account is not dependent on any particular environment structure. To demonstrate this, I present four different network types: Erdös-Renyi random graphs, scale-free networks, small-world networks, and scale-free small-world networks. Erdös-Renyi random graphs are a well-understood and useful benchmark, with edges placed between nodes selected uniformly and at random (Erdös & Rényi, 1959). Scale-free networks reflect the long-tailed rich-get-richer phenomenology common to many real-world structures such as scientific collaboration networks, the world wide web, and semantic networks (Goh, Kahng, & Kim, 2001; Johns & Jones, 2010; Kello et al., 2010; Steyvers & Tenenbaum, 2005). Small-world networks reflect local clustering combined with short path lengths, which are also common features of semantic networks (Cancho & Solé, 2001; De Deyne & Storms, 2008; De Deyne, Verheyen, Perfors, & Navarro, 2015; Kovács, Bóta, Hajdu, & Krész, 2021; Steyvers & Tenenbaum, 2005).

The standard versions of the above network types are often unweighted. To reflect the differential learnability of associations, I develop a rank-based framework which produces weighted random networks based on each of the above network types. This method is a variation of a fitness-based network model that starts with all nodes present and adds edges over time by choosing nodes according to a specific probability distribution (Hills, 2025; Menczer, Fortunato, & Davis, 2020). Each concept is assigned a rank, r, from 1 to the number of concepts, n. Concepts are randomly assigned to M communities. A concept, i, is sampled with probability $P_i \propto r_i^{-\gamma}$. A node's probability of selection is analogous to the frequency with which the concept is encountered. Then a second concept is chosen with probability $P_j \propto r_j^{-\gamma}\left[p_c\delta_{c_i,c_j} + (1-p_c)(1-\delta_{c_i,c_j})\right]$, where δ_{c_i,c_j} is the Kronecker delta function equal to 1 when both nodes are in the same community, $c_i = c_j$, and otherwise 0. Preference for connecting to nodes in the same community is denoted by $p_c > 0.5$. After each sampling, 1 is added to the edge weight between i and j. I explain each component further below.

When $M = 1$ all nodes are in the same community and sampling reduces to $P_i \propto r_i^{-\gamma}$ for both nodes. When $M = 1$ and $\gamma = 0$ the resulting network is a weighted Erdös-Renyi random graph, in which all node pairs have an equal chance of increasing their edge weight. When the community size M is greater than 1, this adapts the approach for developing multi-community networks described by Girvan and Newman (2002): Edges within communities are more likely than edges between communities, with probability p_c and $1 - p_c$, respectively. When $M > 1$ and $\gamma = 0$, the network is a weighted approximation of the community formation model provided by Girvan and Newman (2002). As γ grows larger than 0 this approaches a scale-free distribution and increases the probability of selection of nodes of smaller r. Scale-free distributions are variable in practice and variously defined (Broido & Clauset, 2019). In the present formulation, γ is set to 1 to prevent the over-production of isolates with high r.

Our goal here is not to replicate the exact nature of experience, which can differ dramatically even among young children (Hart & Risley, 1995), but to demonstrate that the effects of enrichment produce predictable outcomes on behavior across a broad range of environments. To demonstrate that, the Supplementary Material shows that the qualitative pattern of results presented here is not altered by substantially increasing or decreasing the number of concepts or associations sampled to construct the environment.

For the purposes of computational tractability and to demonstrate learning towards network saturation, the environments each have 500 concepts. From these concepts, 2000 pairs of concepts are sampled with replacement to form the weighted relationships between them, with the weight corresponding to the number of times the pair is chosen across the 2000 samples. The relationships in the environmental network are therefore represented by weighted and undirected edges.

Figure 2 provides an example of the four environmental network types and for each network the distribution of its node strengths (the sum of each node's edge weights) and degree. This further corroborates the power-law nature of the scale-free networks for both degree and strength. It also shows that small-world networks share similar strength distributions with their corresponding non-small-world counterparts.

Representation

Cognitive representations are built by sampling from and then learning about relationships in the environment. For learning, I use the prediction error framework set out

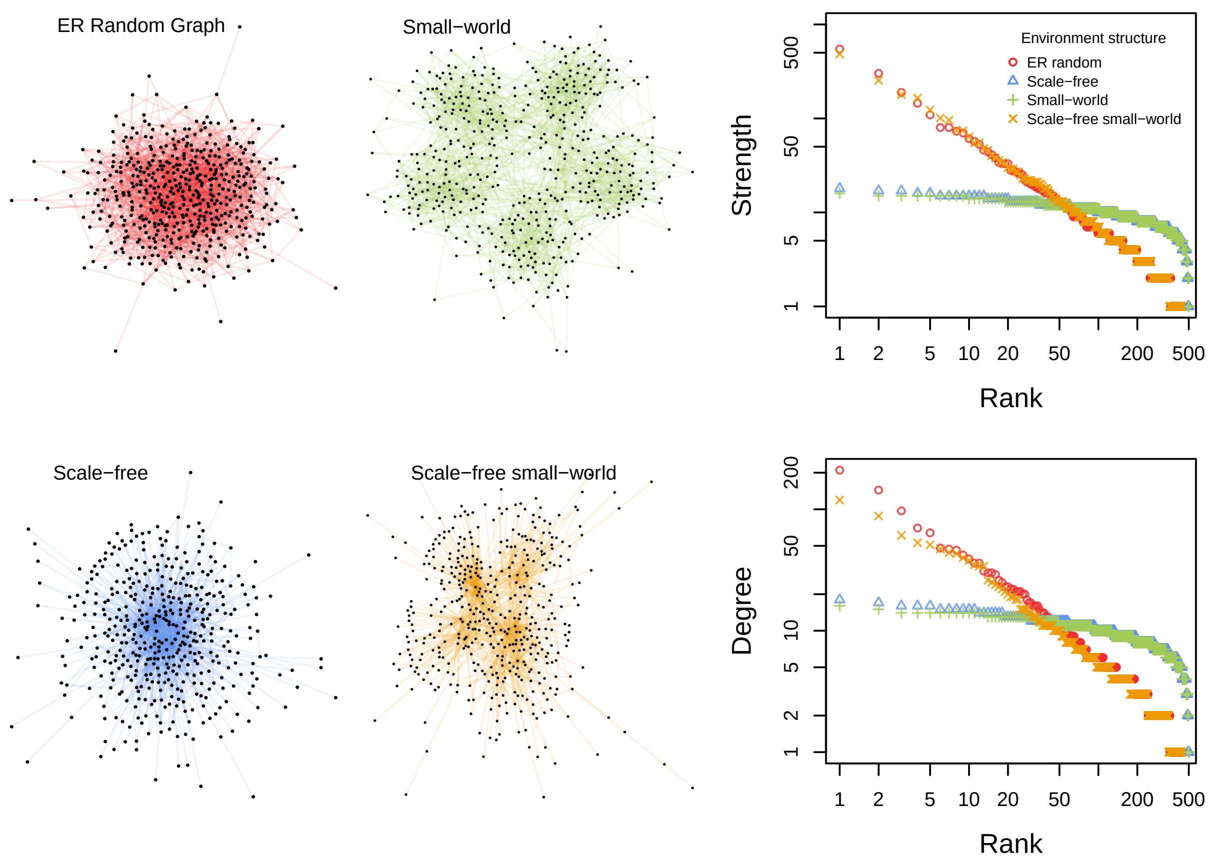

Figure 2. The structure of the environment from which learning takes place. Four weighted network types are shown: An Erdös-Renyi random graph ($\gamma = 0$ and $M = 1$), a small-world network ($\gamma = 0$, $M = 5$, and $p_c = .96$), a scale-free network ($\gamma = 1$ and $M = 1$), and a scale-free small-world ($\gamma = 1$, $M = 5$, and $p_c = .96$). Each network has 500 concepts (nodes) and the weighted edges between them are the result of repeatedly sampling 2000 pairs of nodes and adding 1 to the edge weight between the pairs each time as described in the main text.

in the Rescorla-Wagner model (Rescorla & Wagner, 1972). The choice of the Rescorla-Wagner model follows the substantial evidence for learning as a process of minimizing prediction error, which is a fundamental assumption among models of reinforcement learning (Dayan & Abbott, 2005; Hoppe, Hendriks, Ramscar, & Rij, 2022; McClelland & Rumelhart, 1981; Sutton & Barto, 2018). The Rescorla-Wagner model captures this phenomenology—including associative learning, blocking, inhibition, and extinction—and it is a model on which many subsequent models have been based (e.g., Sutton & Barto, 1981). Though it is not without limitations (Miller, Barnet, & Grahame, 1995; Yau & McNally, 2023), these limitations are largely irrelevant here and the Supplementary Material shows that removing one point of controversy (the context cue) does not alter the results. I therefore use the model to capture the generic prediction-error process inherent in the Rescorla-Wagner design and in recognition of its predictive utility (Roesch, Esber, Li, Daw, & Schoenbaum, 2012; Soto, Vogel, Uribe-Bahamonde, & Perez, 2023; Trimmer, McNamara, Houston, & Marshall, 2012).

Formally, the Rescorla-Wagner model minimizes the prediction error between the values of an observed outcome, j, and a cue predictive of that outcome, i. The value of the outcome is λ_j and the value for that outcome as predicted by the cue is $V_{i\to j}$. The prediction error is the difference between them, $(\lambda_j - V_{i\to j})$, and it is minimized following each learning event according to the following rule:

$$\Delta V_{i\to j} = \alpha_i \beta_j (\lambda_j - V_{i\to j})$$

The parameter α_i corresponds to cue salience (some cues are easier to learn about than others) and β_j to the learning rate for outcomes (some outcomes are easier to learn about than others). Both α and β values are confined to values between 0 and 1. After learning at time t, the updated cue value is

$$V_{i\to j,t+1} = V_{i\to j,t} + \Delta V_{i\to j,t}$$

Thus, with repeated experience, $V_{i\to j}$ will approach the observed value λ_j.

The cognitive representation is formed by applying the Rescorla-Wagner model to the environment in the following way. Each learning event randomly samples a relationship (i.e., edge) from the environment in proportion to its weight (the number of times it was sampled during the environment's construction). Of the two concepts associated with the relationship, one is always assigned as the cue and the other as the outcome. The representation is then updated according to the Rescorla-Wagner model, and the learned association is taken to represent an undirected weight of association. The Supplementary Material shows that directed networks produce the same qualitative patterns as shown below.

We let all outcomes be equivalent and associated with $\alpha = 1$ and $\lambda = 1$. To demonstrate that the qualitative results are not dependent on the precise learning values, β is varied from .01 to .1 here, and over its entire range in the Supplementary Material, consistently reproducing the pattern described here. Since α and β are a product and not assigned to individual concepts, changing β is equivalent to changing the product of the two.

To simulate development, learning occurs over 1000 learning trials. This is split into 4 epochs with 250 learning trials each, allowing us to track the developmental pattern across successive ages. The precise number of learning trials per epoch is unrelated to the qualitative pattern of results (see Supplementary Materials). For the analyses that follow, negative edges are removed as they have no intuitive interpretation for entropy, similarity, or free associations. Additional discussion of negative edges and their irrelevance (demonstrated by removing the context cue which creates them) can be found in the Supplementary Material.

Figure 3 provides an example of learning representations over the four epochs for a scale-free small-world environment with $\beta = .01$. This moves from sparse connectivity on the left, to highly interconnected "hairballs" on the right, a pattern common to all environments. Table 1 quantifies the statistical properties for each of the four environments, showing the mean values for 1000 simulations and their resulting cognitive representations after learning across the simulated lifespan. More learning leads to increasing interconnectivity in the representation. The total number of nodes with associations increases, corresponding to a rising functional vocabulary. The total number of associations increases, as measured by the total number of edges and mean degree. The strength of associations increases, as measured by the sum of a concept's weighted edges with other concepts. The average shortest path length falls, meaning that random concepts share shorter pathways through the network. And the clustering coefficient increases, meaning that the neighbors of concepts tend to become more connected among themselves. Finally, the modularity—a measure of the discriminability of groups within the networks—falls with age, but is highest for the small-world networks, consistent with their community structure. These statistical patterns provide objective measures of cognitive network enrichment.

Behavior

The first two observations associated with cognitive aging we aim to explain are a rising entropy of associations and a reduction in pairwise similarity judgments. Each of these is recovered from the learned cognitive representation as described below. In all cases, behavioral measures are applied within the large subnetwork of interconnected and mutually reachable concepts (i.e., the giant component) making up the majority of words in adult semantic networks and on which prior work has focused (De Deyne, Navarro, Perfors, Brysbaert, & Storms, 2019; e.g., Steyvers & Tenenbaum, 2005).

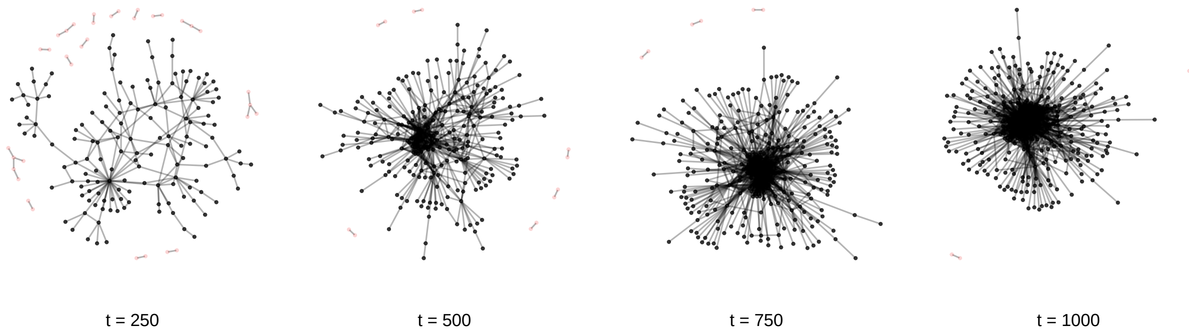

Figure 3. Example of the growing cognitive representation resulting from experience with the scale-free small-world environment shown in Figure 2. Learning uses the Rescorla-Wagner model with $\beta = 0.01$. Cognitive representation growth from other environments looks visually similar. Statistics for these environments are provided in Table 1. Training occurs in 250 event epochs, with edges from the environment sampled in proportion to their weight. Nodes represent individual concepts and edges represent learned associations.

Rising entropy. Following past work (Dubossarsky et al., 2017; Fradkin & Eldar, 2022; Stella, 2020), I use information entropy to quantify the surprisingness of associative responses given the cue. If a cue has only one or a few strongly weighted associations, any given associate will be less surprising than if it has many equally weighted associations. Because each concept is connected to other concepts in the representation by a set of weighted edges, we can compute the entropy for every cue in the network representation as follows:

$$H = -\sum_{i=1}^{k} p_i log(p_i)$$

Here, p_i is the proportion of the weight w_i along edge i with respect to all k edges for that cue. That is, $p_i = \frac{w_i}{\sum_k w_k}$. The entropy reported below is the mean entropy across all concepts which are shared across all epochs. This would be the entropy that would be recorded if one or more individuals produced many associations for each cue, with each association produced in proportion to its edge weight.

Falling similarity. To simulate similarity judgments, we create a measure of co-activation between pairs of cues using spreading activation. Spreading activation and its close cousin random walks are a commonly used process for evaluating co-activation on cognitive representations (Collins & Loftus, 1975; McClelland & Rumelhart, 1981; Siew, 2019; Vitevitch & Mullin, 2021). It also has a rich history in the evaluation of similarity (Goldstone, 1994; Kumar, 2021; Larkey & Markman, 2005; Rotaru, Vigliocco, & Frank, 2018) including predicting human similarity judgments (De Deyne, Navarro, Perfors, & Storms, 2016). To model this, we allow spreading activation to leave one node in the pair and measure activation at the other node, $A_{j \to k}$. This allows us to measure the extent to which one word co-activates the other. Doing this for both cues, we take similarity as the summed co-activation.

$$S = A_{j \to k} + A_{k \to j}$$

We measure this similarity for a random selection of 20 concept pairs in the representation. This only uses concepts learned during the first epoch of learning, ensuring that all concepts are included in the vocabulary for each epoch. Simulations use the `spreadr` function from Siew (2019). This simulates 100 units of activation released from the cue concept, j, and divides its activation along each associative edge in proportion to their weights at each time step. The maximum activation at the target concept over 10 time steps is recorded as $A_{j \to k}$. Then the cue and target are swapped and the simulation is repeated to capture $A_{k \to j}$. The retention, suppress, and decay parameters in `spreadr` are all set to 0, which means all activation remains in the network and is diffused from each node in entirety at each time point. As the parameter values for retention, suppress, and decay increase, spreading activation is dampened, but the results remain qualitatively similar up to the point when the effects are sufficiently dampened that no cross-activation occurs.

Results for entropy and similarity. The results of the above computations are shown in Figure 4, with 100 simulations of the environment, learning, and behavior for each set of parameter values. These are shown alongside the data from Dubossarsky et al. (2017) and Wulff et al. (2019). The results consistently show that entropy rises and similarity falls with increased enrichment, following the qualitative patterns in the observed data.

Faster learning rates have a faster rise in entropy and a faster decline in similarity, consistent with the proposal that network enrichment drives these effects. The correspondence between entropy and similarity suggest they are both influenced by similar underlying structure, which is more rapidly felt when learning is faster. In addition, scale-free networks have a higher entropy and a lower similarity than their non-scale-free counterparts. Community structure ($M > 1$), on the other hand, appears to have a limited influence on the effects of enrichment. As further demonstrated in the Supplementary Materials, this qualitative fit to observation is robust to a wide variety of assumptions about environment size, density, learning algorithms, and duration of learning. From this we can conclude that the enrichment account provides a qualitatively constrained set of outcomes.

Researchers often attempt to bolster the sufficiency of their model by quantitatively fitting it to the data. This is possible with the enrichment account, but it requires additional assumptions—the size and structure of the environment, learning rates and amounts across the lifespan—which are likely to be overspecified. In these and other situations, Roberts and Pashler (2000) persuasively argue that model fitting can be misleading. In particular, model fitting can be misleading for three reasons: a) models can

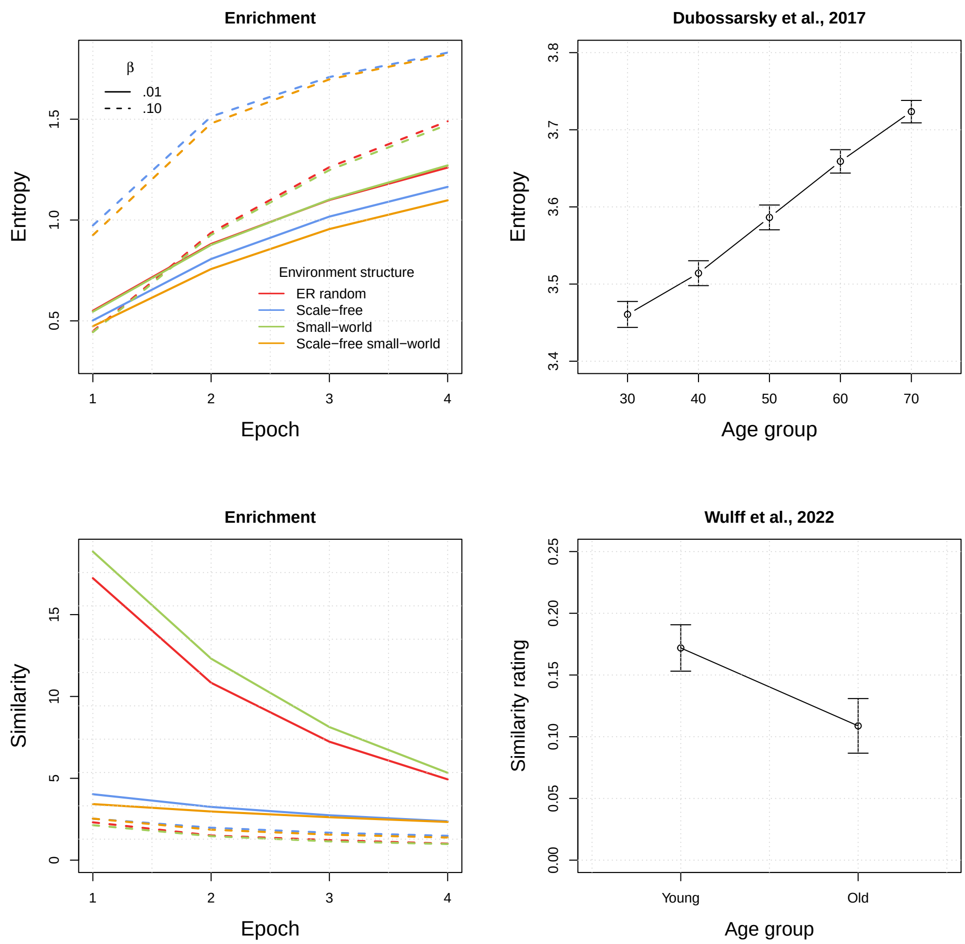

Figure 4. The outcome of the representational enrichment model for entropy and similarity measures compared against the Dubossarsky et al., 2017 and Wulff et al., 2022 data. Results represent 100 simulations of each environment type, and entropy and similarity computed from the learned cognitive representations with β equal to .01 or .10. Entropy is based on the distribution of edge weights. Similarity is based on the similarity equation for cross-activation via spreading activation. Using the free association data described in Table 1 of Dubossarsky et al. (2017) for ages 30 and above, entropy is computed as described for the enrichment model. The Wulff et al. (2022) data is based on 2268 similarity ratings for each of 36 old and 36 young individuals. Data are means and standard error after averaging first within participants and then over participants in an age group.

be unconstrained, b) theory assumptions may be inconsistent with prior theory and evidence, and c) other models with different assumptions could fit the data equally well. The cognitive network enrichment account safely handles the first two challenges: a) over a range of parameter values, it always predicts the qualitative pattern in the data—it cannot fit any pattern—and b) it is based on applying established theories of learning and memory across the lifespan. What about alternative models? Could degradation explain the observed effects?

Degradation

All other things being equal, older adults follow the general pattern found across the lifespan, they tend to remember things for which they have more experience (e.g., D. M. Burke & MacKay, 1997). This has a straightforward mapping onto degradation: Degrade associations by removing the weakest relationships in the cognitive representation first (e.g., Borge-Holthoefer et al., 2011). I implement this by ranking all edges in order of weight from weakest to strongest, and then removing all edges in the bottom 20%, 40%, 60% and 80%. I also simulate random edge removal. Though this does not have a known correspondence to memory loss, it provides a useful counterpoint. This removes associations without reference to their frequency of past experience. This is simulated by sampling 20%, 40%, 60% and 80% of the edges at random and then removing them.

Because degradation requires prior learning, I start with the fully learned network that is the end state of learning in Figure 4. That is, starting with the epoch 4 representation I remove various proportions of edges as a proxy for degradation and compare these alongside an additional epoch of enrichment. This produces a qualitative model competition in the fifth epoch, one in which we can directly compare enrichment alongside various kinds and intensities of degradation. The simulation is repeated 100 times for each environment type, starting with 500 concepts and 2000 associations. There are 250 learning events per epoch, using $\beta = 0.1$. Values for similarity and entropy are computed as above.

As Figure 5 shows, degradation is inconsistent with enrichment and it is inconsistent with the observed data. Both weakest first and random edge removal produce falling entropy and rising similarity. These effects become more severe as degradation increases. Degradation, as formally described here, is not sufficient, nor is it necessary.

Free association networks across the lifespan

Does enrichment explain the sparsening of free association networks found by Dubossarsky et al. (2017), Jeong and Hills (2024), and Zortea et al. (2014), which was also shadowed in the memory search networks of Wulff, Hills, et al. (2022)? In all cases, age was marked by decreasing network degree, increasing average shortest path length, and decreasing clustering coefficient. Paradoxically, this is the opposite pattern to that observed for the enriched representations shown in Table 1: more learning leads representations to rise in degree, fall in average shortest path length, and rise in clustering coefficient. Does retrieval from these cognitive representations reproduce the observed pattern for free associations? I test this here by simulating free association retrieval from the enriched representations and then building networks from those free associations following the procedure used in Dubossarsky et al. (2017).

For each cognitive representation, we simulate 5 participants who retrieve three associates from each of 30 randomly chosen cues. The cues are sampled from the subset of concepts that have at least three associates in each epoch. The three associates are sampled without replacement for each participant and each associate is sampled in proportion to the positive associative strength with its cue encoded in the cognitive representation. This produces a cue-by-associate matrix, with each cell indicating the number of times each associate was produced in response to each cue.

To control for varying numbers of associates produced across the lifespan,

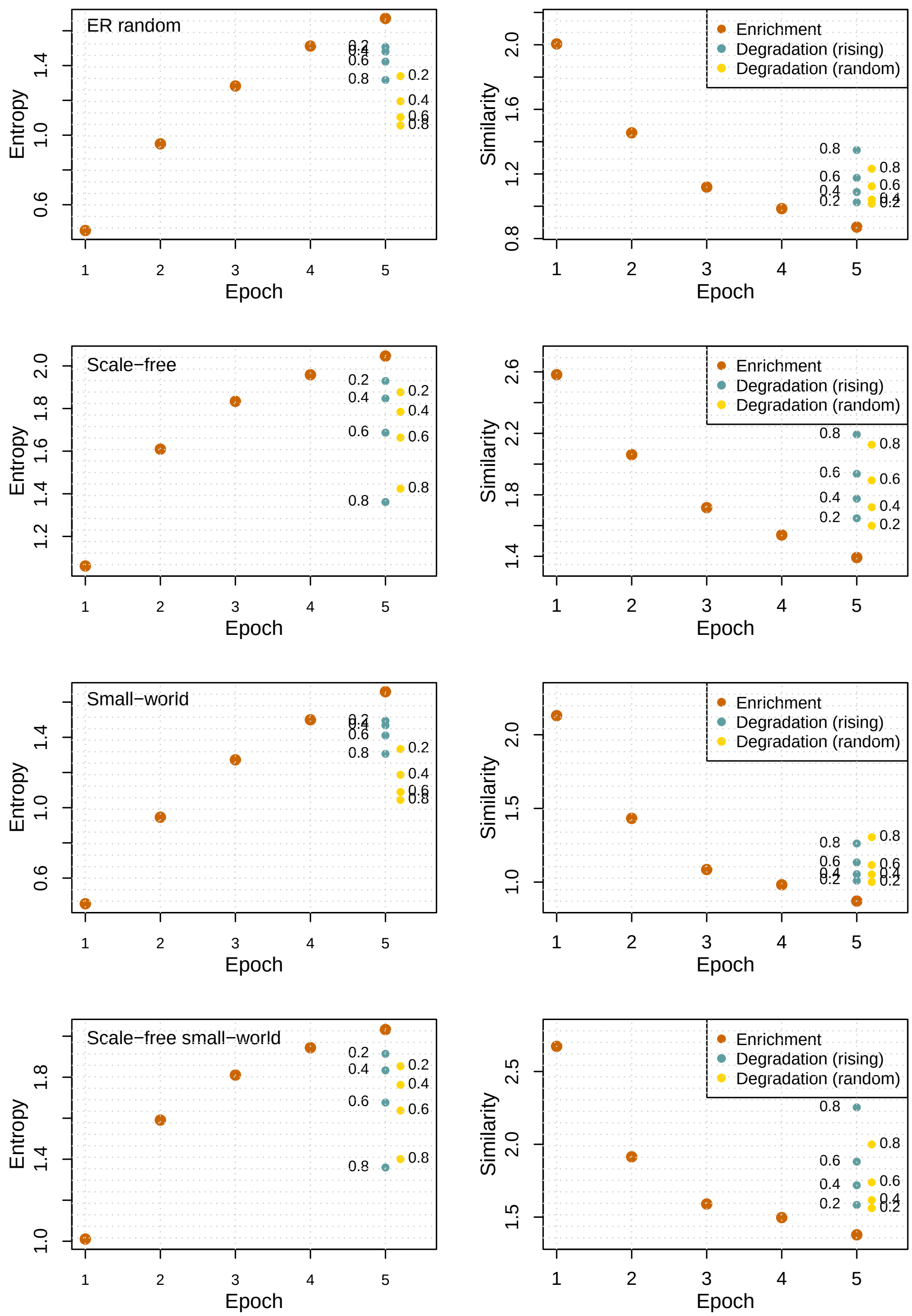

Figure 5. Comparisons of enrichment alongside various kinds and intensities of degradation. Degradation is created either through a 'rising' threshold (weakest first) or 'random' edge removal starting from Epoch 4. The proportion of degradation is indicated alongside each point.

Dubossarsky et al. (2017) converted the cue by association network into networks of equivalent size by producing a cue-by-cue network. Weighted edges in the cue-by-cue network were computed as the count of associates from one cue to the associates of the other cue, as follows,

$$w_{i,j} = \sum_{a \in P} \frac{w_{i,a}}{N_a - 1}$$

with $w_{i,j}$ being the weighted directed edge from cue i to cue j, P is the set of shared associates, $w_{i,a}$ is the number of times cue i produced associate a in P, and N_a is the number of cue words that produced a. Thus, if two cues each produce one common associate that is not produced by other cues, then $w_{i,j} = w_{j,i} = 1$. After transforming the cue-by-associate network into the cue-by-cue network, Dubossarsky et al. (2017) removed edges with a weight lower than 1. As the networks are smaller here, no threshold is used in the first instance, but I vary this and other parameters below.

From these networks, degree, average shortest path length, and clustering coefficient are computed. Dubossarsky et al. (2017) computed degree following Opsahl, Agneessens, and Skvoretz (2010) as a combination of degree, k, and strength, w (the sum of the weights), using $k = \sqrt{k_i} \times \sqrt{w_i}$. Average shortest path length was computed by transforming weights into distances by dividing by the average weight of the network, again following Opsahl et al. (2010). Clustering coefficient was computed for weighted networks following the method described in Barrat, Barthelemy, Pastor-Satorras, and Vespignani (2004). This entire process is repeated for each environment type 100 times, with learning and then free associate retrieval as described above.

The results, shown in Figure 6, follow the empirically observed pattern, with more learning leading to declining degree, rising average shortest path length, and falling clustering coefficient. This is true across all four network types, suggesting this effect is not a property of the environment. A clear take-home message is that the apparent sparsening of the free association network can, counter-intuitively, be driven by enrichment of the cognitive representation.

However, there is more to be learned here. The method for collecting data from participants and constructing networks is diverse and results can be sensitive to these details. All previous work uses varying numbers of participants, cues, and thresholds for including edges in the network, with some diversity of results (Dubossarsky et al., 2017; Jeong & Hills, 2024; Wulff, De Deyne, et al., 2022; Zortea et al., 2014). This is to be expected. If older adults produce higher entropy responses, older participants will generate less overlap per response, and this will produce weaker associations in the free association network. Application of a threshold—a common practice with free association data—also removes weak edges and makes older adults' networks look more sparse. Keeping data collection parameters similar for older and younger adults, such as participant numbers, cues, or associations, is effectively a hidden threshold because higher entropy responses will require more data to capture overlap in the distribution.

With more participants, older adults begin to overlap across a broader set of associations, and cue degree increases. Indeed, networks with higher degree, lower average shortest path length, and higher clustering coefficients arise as more participants report more associates for more cues. This is shown in the parameter exploration in Figure 7, which provides an example using the scale-free small-world environment. This is a representative sample of the more complete exploration, which also includes degradation, shown in the Supplementary Material. What these explorations show is that for all environment types, there is always a parameter range for which enrichment (and degradation) captures the empirical evidence for sparsening of free association networks with age.

The versatility of the models in capturing various patterns of free association data should be unsettling. Given that we already know degradation does not reproduce the

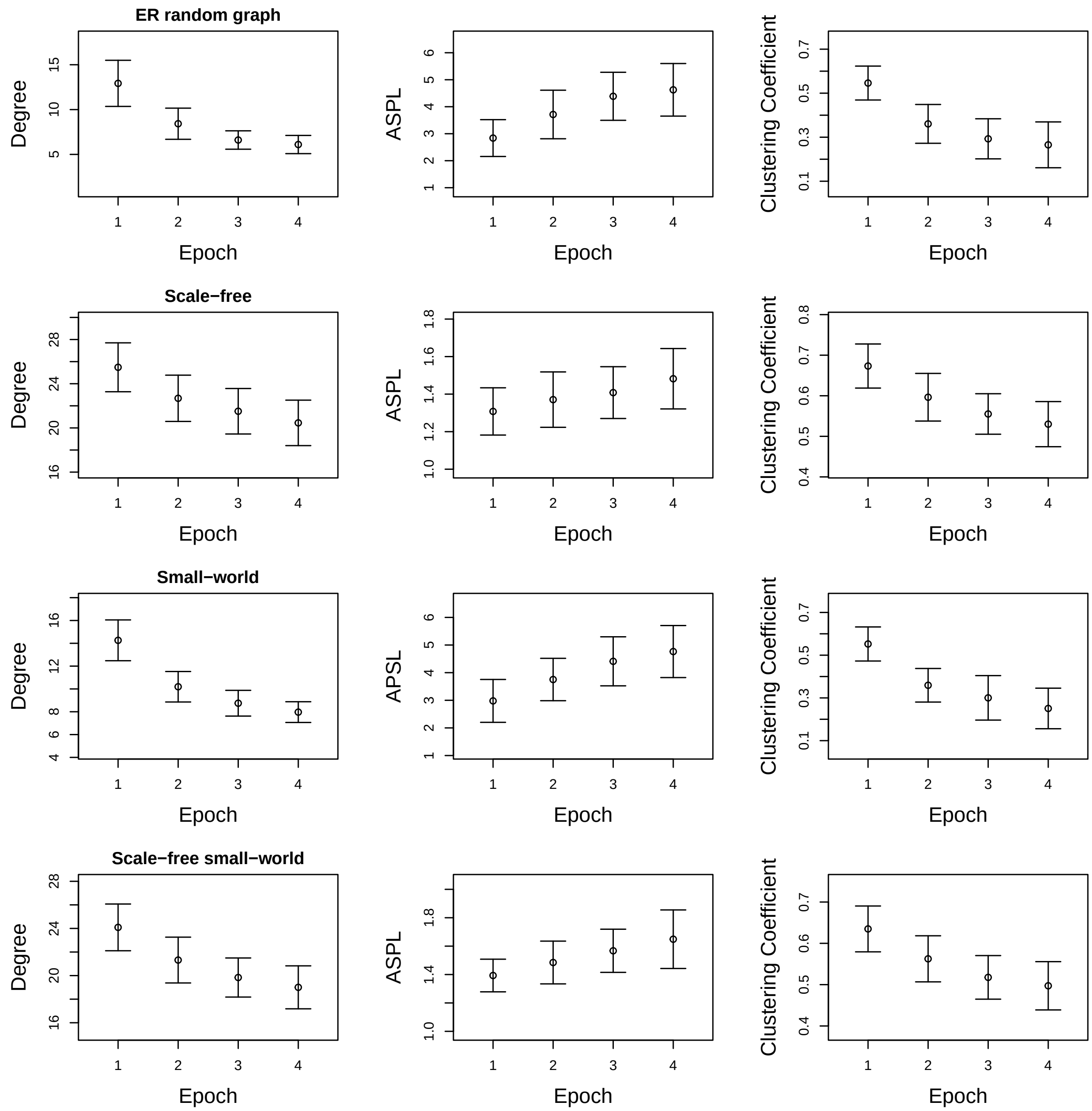

Figure 6. Free association networks as a consequence of enrichment, showing the degree, average shortest path length (ASPL), and clustering coefficient with increasing age. Simulations were repeated 100 times for each environment type, with four training epochs of 250 learning events each ($\beta = .1$). Representations were then each sampled from by 5 simulated participants who each retrieved 3 associates for each of 30 cues with probability proportional to the associative strengths in the learned representation. Free association network construction and measures are as described in the text. Error bars indicate one standard deviation.

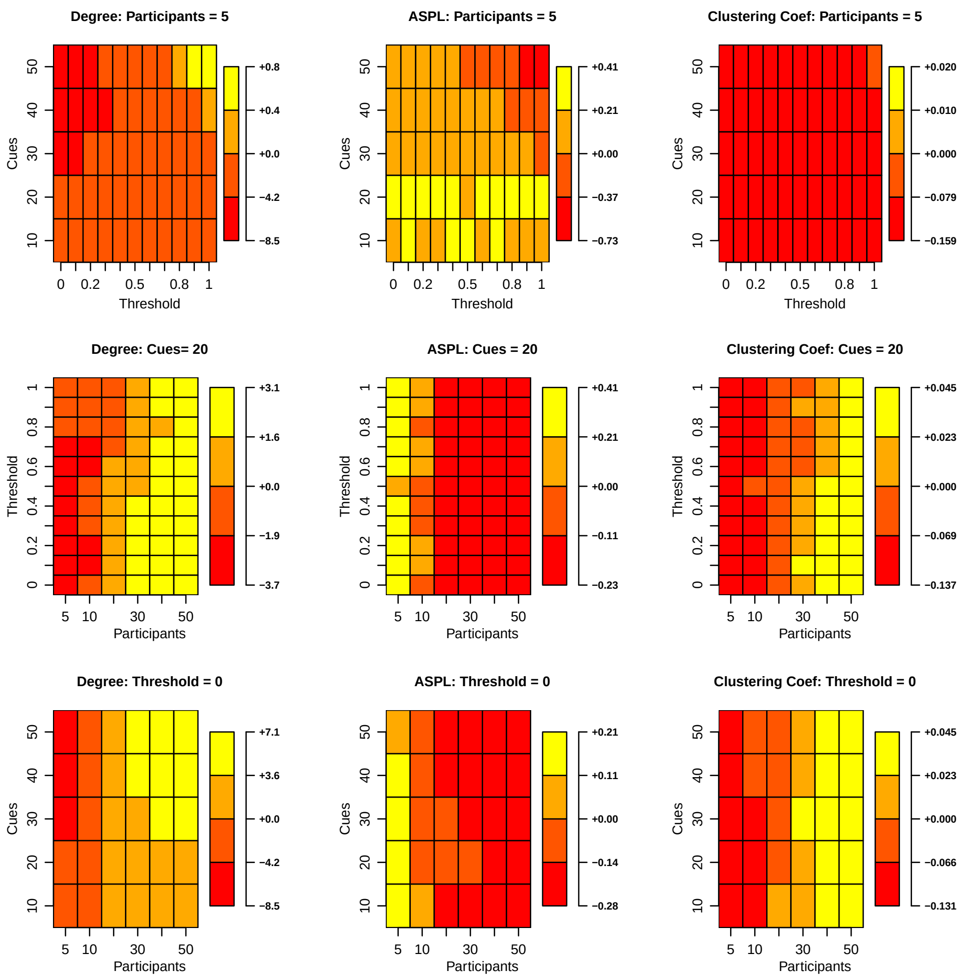

Figure 7. Parameter exploration varying number of participants, cues, and threshold for the effects of enrichment on free association network structure. Simulations were repeated for 100 different scale-free small-world environments ($\gamma = 1$ and $M = 5$) with four training epochs each of 250 learning events each ($\beta = .1$). Representations were then each sampled from by the indicated number of participants who each retrieved 3 associates for each of the cues with probability proportional to the associative strengths. Cue-by-cue networks were then constructed as above. The heatmap indicates the difference between Epoch 4 and Epoch 1, with numbers greater than 0 indicating increasing values across the lifespan, shown in yellow, and declining values shown in red.

entropy or similarity results, degradation's success here is merely warning. That both models can produce the same apparent sparsening further compounds the challenge of inferring the representation from the behavior and the mistake of conflating the two.

However, there are other clues supporting enrichment in this case. Older adults produce more unique and unrelated associations than younger adults given the same number of cues (Dubossarsky et al., 2017; Jeong & Hills, 2024; Zortea et al., 2014). In the limit, if older adults chose from all possible associates, a sufficient amount of data would produce a fully connected free association network—everything would be connected to everything. Collecting data from too few participants or across too few cues would however produce weaker associations than for younger more predictable participants. This is problematic as young children may also produce few associations because they know fewer words. If fundamentally different cognitive representations can produce the same behavioral patterns depending on methodological details, then further theory is needed to help identify the limits of our inference and how best to overcome them.

Discussion

Cognitive science, in the words of Johnson-Laird (1980) "needs theories that both cohere and correspond to the facts" (p. 73). The cognitive network enrichment account does both in the sense that it relies on well-established principles of learning and cognitive retrieval, models from environment to behavior, and produces observed cognitive effects commonly associated with aging across the lifespan. As cognitive representations become enriched and more densely interconnected through lifelong experience, the behaviors resulting from them lead to rising entropy in associations and falling judgments of similarity. An account based on degradation, on the other hand, fails to capture these results.

The enrichment account, like any model, helps put guardrails on our ignorance. It does this by showing how many of the behaviors that we might have intuitively interpreted as evidence of degradation are straightforward outcomes of learning. An enrichment of representational associations leads to higher behavioral entropy in divergent tasks and a greater diffusion of activation away from potential targets in convergent tasks. Indeed, we know enrichment must be occurring in association with rising crystallized intelligence. Thus the enrichment account is not only parsimonious, its effects *should* be predicted, even if degradation is influencing them in some as yet unknown way.

In addition to the criteria of constrained predictions, coherence with prior theory, and outperforming alternatives, the enrichment account satisfies a further requirement set out for models by Roberts and Pashler (2000): the outcome is surprising. If degradation-based cognitive decline is the a priori explanation for apparent sparsening of the mental lexicon, age-related slowing, and entropy of responses, then the observation that these are all explained by assuming people continue to learn as they always have, without invoking degradation at all, is surprising. Degradation would reverse the effect (see also Borge-Holthoefer et al., 2011). It is still more surprising because the representation that gives rise to apparent sparsening is not itself sparsening.

These findings also help sharpen our intuition surrounding representation and behavior. Though only enrichment can explain the entropy and similarity effects, both enrichment and degradation can explain age-related change in the structure of free association networks. The latter observation should make researchers wary of assuming that any behavior, including free association, is a direct readout of the underlying representation. In the case of entropy and similarity judgments, inferring representation directly from behavior without formally taking into account the processes needed to generate that behavior leads to process inflation: we infer an additional process—degradation—where it is unnecessary.

Many further challenges arise from these results. While the behavioral patterns are one kind of data, the patterns of biological and neural change outlined in the introduction are data as well. Any complete model of cognitive aging must take these facts into account, bridging the gap between cognition and its biological substrate. Ideally, we want to understand how these changes in neural architecture interact with enrichment to produce the behavioral variation that we see. One provocative interpretation is that degradation accounts may get the causation backwards. Which is to say, if life experience gives rise to some of the primary markers of healthy age-related cognitive change, it may also contribute to developmental patterns in the brain we associate with that change. Research focused on neural enrichment may help better understand this potential pathway.

The enrichment account also invites us to consider what aspects of development are non-stationary. Here I assumed the environment, encoding, and retrieval were all fixed across the lifespan. But learning strategies may change in relation to age or content (Luef, 2022; Stella, Beckage, & Brede, 2017). Childhood may have special developmental features of its own. For example, early vocabulary acquisition is driven by a variety of processes including semantics, phonology, social pragmatics, and perceptual features (Ciaglia, Stella, & Kennington, 2023; Fourtassi, Bian, & Frank, 2020; Siew & Vitevitch, 2020; Yu, Suanda, & Smith, 2019). Developmental patterns that extend into young adulthood (e.g., because of shared education) may change as individuals age further, separating early from late life (compare Dubossarsky et al., 2017; Jeong & Hills, 2024). Investigating how these additional factors impact lifelong development may help us better understand conceptual development and the nuances that may make some aspects of cognition more prone to enrichment effects than others.

Another non-stationary factor is change in the conceptual environment. Cultural evolution that occurs during a lifetime may produce new concepts and artifacts or redistribute associations among concepts that already exist (Hills & Adelman, 2015; Li, Engelthaler, Siew, & Hills, 2019; Li & Siew, 2022). Similarly, familiarity with concepts may make novel higher or lower order concepts available (Goldstone, Rogosky, Pevtzow, & Blair, 2017; Kemp & Tenenbaum, 2008). These non-stationarities provide new opportunities for learning that may dilute or further enrich existing conceptual relationships. Inspired by Brysbaert et al. (2016), Hills (2025) demonstrated that a minimalist textbook exploration of environmental growth according to Herdan-Heaps law (adding new concepts to the environment with additional experience) could reproduce the entropy and similarity effects described here. The present work shows that such environmental growth is not necessary, and extends the effects of enrichment to novel environments and behaviors. However, there are many other ways in which environments could change, including differences in where new words arise, the effects of specialization, and the influence of changes in association weights or environmental ranking. Taking this seriously requires a multi-disciplinary modeling approach that treats the conceptual environment as a moving target.

Despite these challenges for future work, the fundamental mechanism of the enrichment account has experimental support. The fan effect demonstrates that learning many relationships with a target concept reduces the speed of accessing any one of those relationships (Anderson & Reder, 1999). The fan effect is also amplified in older adults (Cohen, 1990; Gerard, Zacks, Hasher, & Radvansky, 1991). This diffusion of activation is a general and well-understood outcome of spreading activation and random walker models (Abbott, Austerweil, & Griffiths, 2015; Siew, 2019). Activation follows pathways and the more pathways there are the less targeted the activation. In addition, one can see the enrichment effect in natural experiments: Ramscar, Sun, Hendrix, and Baayen (2017) showed that individuals with more language experience (native speakers) were more impaired in paired-associative learning in that language than age-matched individuals with less experience (second language learners). This suggests the effect is not about age, but enrichment.

The effects of enrichment should should also produce slowing. If we take the substantial empirical support that human judgment and decision making are underpinned by evidence accumulation models (Brown & Heathcote, 2008; Zhu, Sundh, Spicer, Chater,

& Sanborn, 2024), then higher entropy predicts longer reaction times. This can be easily demonstrated on the back of an envelope. Suppose we want to retrieve B from memory following cueing with A. If concept A is equally connected with concepts B and C, the negative binomial gives the expected number of samples needed to reach a threshold number of r samples of B as $E = r/p$. If it takes 10 successes to reach the threshold, then it would take 20 samples on average for B to reach the threshold. In contrast, if the probabilities were 0.7 and 0.3, for B and C respectively, then it would take ≈ 14.3 samples, leading to a faster response time. Thus, all other things being equal, higher entropy in the representational structure leads to slower processing.

Enrichment will also degrade fluid intelligence in other ways. The two most well-supported explanations for working memory constraints are based on resource limitations and interference (Oberauer et al., 2016), effects not limited to language (Oberauer & Lin, 2024). Interference is likely to arise because, as Amer et al. (2022) points out, older adults suffer from retrieval of more information and less relevant information. Because external cues gain their relevance via their encoded associations in memory (Easdale, Le Pelley, & Beesley, 2019), greater entropy of associations will make targeting appropriate external stimuli more challenging. This also explains the greater impact of external clutter on processing speed in older adults (Amer et al., 2022; McCarley, Yamani, Kramer, & Mounts, 2012). This was indirectly predicted by Hasher and Zacks (1988), who noted that, "If there is an age-related increase in the importance of one's personal values and experiences along with an age-related increase in the tendency to apply these concerns to a wider range of information, more information. . . is likely to enter working memory." Here, we might only add that the increase may be due to greater internal competition as a result of enrichment of representational targets.

Enrichment may also explain the appearance of impaired inhibition in older adults: the activation strengths associated with extraneous and potentially irrelevant information are relatively more competitive in individuals with more enriched representations. This makes older individuals more susceptible to clutter driven by both internal and external distractions. Many aging theorists explain the influence of these distractors as deficits in inhibition (e.g., Stoltzfus, Hasher, & Zacks, 1996). This focuses on the *process* as the target of age-related decline. The present work reminds us that the outcomes of this process will be directly affected by representational structure (Hills, 2025). The control processes associated with inhibition may be identical in older and younger adults, but older adults may suffer under greater competition from an enrichment of prior knowledge, making inhibition harder.

Should the results of the enrichment account generalize to tasks that are less based on prior knowledge? Tasks such as Raven's Advanced Progressive Matrices, trail-making tasks, or yet other fluid intelligence tasks are often described as knowledge-free. Ramscar et al. (2014) argues cogently that it is quite difficult to establish that a task is language-free. The bar should be raised still higher for producing tasks that are knowledge-free. For example, in an ostensibly knowledge-free decision task (the leapfrog task), Blanco et al. (2016) showed that older adults relied more on prior strategies than younger adults. This served them well when prior strategies were effective (they outperformed younger individuals), but it served them poorly when those strategies were misleading (younger adults performed better). To the extent that tasks can be free of prior knowledge, the enrichment account predicts they should be less influenced by age. But even Raven's Matrices is influenced by rule learning (Loesche, Wiley, & Hasselhorn, 2015), which would require an encoding independent of prior knowledge to avoid the impact of enrichment. Evidence for that independence is unlikely. A variety of measures of long-term memory correlate with fluid intelligence and working memory tasks (Unsworth, 2019) as do cultural and historical factors commonly associated with the Flynn effect (Brouwers, Van de Vijver, & Van Hemert, 2009). Finally, if we accept more than a century of research showing that learning is itself influenced by prior knowledge then prior knowledge should affect performance on novel tasks. These all suggest that interactions between knowledge representations and knowledge-free tasks should be expected unless rigorously ruled out.

One false challenge to the enrichment account is that individuals with higher education and occupational attainment are less likely to experience late-life cognitive decline (Clouston et al., 2020; Lövdén, Fratiglioni, Glymour, Lindenberger, & Tucker-Drob, 2020; Stern et al., 1994). This is partially a consequence of compensation accorded by skills or strategic repertoires, what is called cognitive reserve (Scarmeas & Stern, 2003). Additionally, there is ample evidence that differences in cognitive skills emerge early in life, prior to education, and therefore lay the foundation for educational attainment later in life (Deary, Whiteman, Starr, Whalley, & Fox, 2004; Lövdén et al., 2020). If differences in processing speed at an early age influence educational attainment at a later age, then slowing later as a consequence of education will be confounded by early processing advantages that facilitated that education. Moreover, these early processing advantages may not be a result of experience at all. IQ is correlated with brain volume (Pietschnig, Gerdesmann, Zeiler, & Voracek, 2022) and is strongly heritable (Peper, Brouwer, Boomsma, Kahn, & Hulshoff Pol, 2007). Furthermore, measures of 'brain age' in late life are well-predicted by indicators present in early life (Vidal-Pineiro et al., 2021). Thus, inferring that educational attainment makes individuals more resilient to late-life cognitive decline assumes a causation that may be absent. The capacity for resilience may derive from the same source as the capacity for attainment.

Finally, the network enrichment effects observed here are not limited to human cognition but are ubiquitous across complex systems. The stability of economic markets, food webs, IT systems, urban infrastructure, neural networks, gene regulatory networks, and many other systems are impacted by network structure (e.g., Strogatz, 2001; Turnbull et al., 2018). In financial networks and food webs, connectivity can dilute resource flows, affecting stability and flexibility (May, 2013). In cultural systems, an increasing capacity to send and receive information creates competition for attention, altering information quality (Hills, 2019; Qiu, FM Oliveira, Sahami Shirazi, Flammini, & Menczer, 2017). Greater connectivity in the brain can facilitate pathology by upseting the balance between integration and segregation (Lord, Stevner, Deco, & Kringelbach, 2017). These systems are each unique, but they often share fundamental topological and resource constraints. Perhaps the aging mind suffers under the same scaling laws as the streets of old Rome (e.g., Bettencourt, 2021). One should not make too much of an analogy, but where constraints are mirrored, phenomenology should follow.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Random walks on semantic networks can resemble optimal foraging. *Psychological Review*, *122*(3), 558–569.

Amer, T., Wynn, J. S., & Hasher, L. (2022). Cluttered memory representations shape cognition in old age. *Trends in Cognitive Sciences*, *26*(3), 255–267.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*(2), 186–197.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging*, *12*(1), 12–21.

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences*, *17*(7), 348–360.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, *101*(11), 3747–3752.

Bettencourt, L. M. (2021). *Introduction to urban science: Evidence and theory of cities as complex systems.*

Blanco, N. J., Love, B. C., Ramscar, M., Otto, A. R., Smayda, K., & Maddox, W. T. (2016). Exploratory decision-making as a function of lifelong experience, not cognitive decline. *Journal of Experimental Psychology: General*, *145*(3), 284–297.

Borge-Holthoefer, J., Moreno, Y., & Arenas, A. (2011). Modeling abnormal priming in Alzheimer's patients with a free association network. *PLoS ONE*, *6*(8), e22651.

Boyle, P. A., Wang, T., Yu, L., Wilson, R. S., Dawe, R., Arfanakis, K., ... Bennett, D. A. (2021). To what degree is late life cognitive decline driven by age-related neuropathologies? *Brain*, *144*(7), 2166–2175.

Broido, A. D., & Clauset, A. (2019). Scale-free networks are rare. *Nature Communications*, *10*(1), 1–10.

Brouwers, S. A., Van de Vijver, F. J., & Van Hemert, D. A. (2009). Variation in raven's progressive matrices scores across time and place. *Learning and Individual Differences*, *19*(3), 330–338.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*(1116), 1–11.

Buchler, N. E., & Reder, L. M. (2007). Modeling age-related memory deficits: A two-parameter solution. *Psychology and Aging*, *22*(1), 104–121.

Burke, D. M., & MacKay, D. G. (1997). Memory, language, and ageing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1363), 1845–1856.

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science*, *13*(1), 21–24.

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*, *7*(1), 30–40.

Cancho, R. F. I., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1482), 2261–2265.

Cao, M., Wang, J.-H., Dai, Z.-J., Cao, X.-Y., Jiang, L.-L., Fan, F.-M., et al.others. (2014).

Topological organization of the human brain functional connectome across the lifespan. *Developmental Cognitive Neuroscience*, *7*, 76–93.

Castro, N., & Siew, C. S. (2020). Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proceedings of the Royal Society A*, *476*(2238), 20190825.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Elsevier.

Ciaglia, F., Stella, M., & Kennington, C. (2023). Investigating preferential acquisition and attachment in early word learning through cognitive, visual and latent multiplex lexical networks. *Physica A: Statistical Mechanics and Its Applications*, *612*, 128468.

Clouston, S. A., Smith, D. M., Mukherjee, S., Zhang, Y., Hou, W., Link, B. G., & Richards, M. (2020). Education and cognitive decline: An integrative analysis of global longitudinal studies of cognitive aging. *The Journals of Gerontology: Series B*, *75*(7), e151–e160.

Cohen, G. (1990). Recognition and retrieval of proper names: Age differences in the fan effect. *European Journal of Cognitive Psychology*, *2*(3), 193–204.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

Cosgrove, A. L., Beaty, R. E., Diaz, M. T., & Kenett, Y. N. (2023). Age differences in semantic network structure: Acquiring knowledge shapes semantic memory. *Psychology and Aging*, *38*(2), 87–102.

Cosgrove, A. L., Kenett, Y. N., Beaty, R. E., & Diaz, M. T. (2021). Quantifying flexibility in thought: The resiliency of semantic networks differs across the lifespan. *Cognition*, *211*, 104631.

Damoiseaux, J. S. (2017). Effects of aging on functional and structural brain connectivity. *Neuroimage*, *160*, 32–40.

Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987–1006.

De Deyne, S., Navarro, D. J., Perfors, A., & Storms, G. (2016). Structure at every scale: A semantic network account of the similarities between unrelated concepts. *Journal of Experimental Psychology: General*, *145*(9), 1228–1254.

De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, *40*(1), 213–231.

De Deyne, S., Verheyen, S, Perfors, A, & Navarro, D. J. (2015). Evidence for widespread thematic structure in the mental lexicon. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *37*. Retrieved from https://escholarship.org/uc/item/2q99m33t

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., ... Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135–152.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the scottish mental surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, *86*(1), 130–147.

Deery, H. A., Di Paolo, R., Moran, C., Egan, G. F., & Jamadar, S. D. (2023). The older adult brain is less modular, more integrated, and less efficient at rest: A systematic review of large-scale resting-state functional brain networks in aging. *Psychophysiology*, *60*(1), e14159.

Desrosiers, G., & Ivison, D. (1986). Paired associate learning: Normative data for differences between high and low associate word pairs. *Journal of Clinical and Experimental Neuropsychology*, *8*(6), 637–642.

Dubossarsky, H., De Deyne, S., & Hills, T. (2017). Quantifying the structure of free association networks across the life span. *Developmental Psychology*, *53*(8), 1560–1570.

Easdale, L. C., Le Pelley, M. E., & Beesley, T. (2019). The onset of uncertainty facilitates the learning of new associations by increasing attention to cues. *Quarterly Journal of Experimental Psychology*, *72*(2), 193–208.

Erdös, R., & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, *6*, 290–297.

Estes, W. (1975). Some targets for mathematical psychology. *Journal of Mathematical Psychology*, *12*(3), 263–282.

Ferreira, L. K., Regina, A. C. B., Kovacevic, N., Martin, M. da G. M., Santos, P. P., Carneiro, C. de G., … Busatto, G. F. (2016). Aging effects on whole-brain functional connectivity in adults free of cognitive and psychiatric disorders. *Cerebral Cortex*, *26*(9), 3851–3865.

Fourtassi, A., Bian, Y., & Frank, M. C. (2020). The growth of children's semantic and phonological networks: Insight from 10 languages. *Cognitive Science*, *44*(7), e12847.

Fradkin, I., & Eldar, E. (2022). Accumulating evidence for myriad alternatives: Modeling the generation of free association. *Psychological Review*, *6*, 1492–1520.

Ge, Y., Grossman, R. I., Babb, J. S., Rabin, M. L., Mannon, L. J., & Kolson, D. L. (2002). Age-related total gray matter and white matter changes in normal adult brain. Part i: Volumetric MR imaging analysis. *American Journal of Neuroradiology*, *23*(8), 1327–1333.

Geerligs, L., Renken, R. J., Saliasi, E., Maurits, N. M., & Lorist, M. M. (2015). A brain-wide study of age-related changes in functional connectivity. *Cerebral Cortex*, *25*(7), 1987–1999.

Gerard, L., Zacks, R. T., Hasher, L., & Radvansky, G. A. (1991). Age deficits in retrieval: The fan effect. *Journal of Gerontology*, *46*(4), P131–P136.

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage*, *51*(3), 943–951.

Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, *99*(12), 7821–7826.

Goh, K.-I., Kahng, B., & Kim, D. (2001). Universal behavior of load distribution in scale-free networks. *Physical Review Letters*, *87*(27), 278701.

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 3–28.

Goldstone, R. L., Rogosky, B. J., Pevtzow, R., & Blair, M. (2017). The construction of perceptual and semantic features during category learning. In *Handbook of categorization in cognitive science* (pp. 851–882). Elsevier.

Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R., & Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *Journal of Cognitive Neuroscience*, *18*(2), 227–241.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Brookes Publishing Co.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, *26*(4), 433–443.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, *22*, 193–224.

Hills, T. (2019). The dark side of information proliferation. *Perspectives on Psychological Science*, *14*(3), 323–330.

Hills, T. (2025). *Behavioral network science: Language, mind, and society*. Cambridge University Press.

Hills, T., & Adelman, J. S. (2015). Recent evolution of learnability in american english from 1800 to 2000. *Cognition*, *143*, 87–92.

Hills, T., & Kenett, Y. N. (2022). Is the mind a network? Maps, vehicles, and skyhooks in cognitive network science. *Topics in Cognitive Science*, *14*(1), 189–208.

Hills, T., & Kenett, Y. N. (2024). An entropy modulation theory of creative exploration. *Psychological Review.*, Advance online publication.

Hills, T., Mata, R., Wilke, A., & Samanez-Larkin, G. R. (2013). Mechanisms of age-related decline in memory search across the adult life span. *Developmental Psychology*, *49*(12), 2396–2404.

Hoffman, P. (2019). Divergent effects of healthy ageing on semantic knowledge and control: Evidence from novel comparisons with semantically impaired patients. *Journal of Neuropsychology*, *13*(3), 462–484.

Hoppe, D. B., Hendriks, P., Ramscar, M., & Rij, J. van. (2022). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*, *54*(5), 2221–2251.

Horn, J. L. (1989). Models of intelligence. In L. R. L. (Ed.), *Intelligence: Measurement, theory, and public policy: Proceedings of a symposium in honor of lloyd g. humphreys* (pp. 29–73). University of Illinois Press.

Jeong, D., & Hills, T. T. (2024). Age-related diversification and specialization in the mental lexicon: Comparing aggregate and individual-level network approaches. *Cognitive Science*, Advanced online publication.

Johns, B. T., Jamieson, R. K., & Jones, M. N. (2023). Scalable cognitive modelling: Putting Simon's (1969) ant back on the beach. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *77*(3), 185–201.

Johns, B. T., & Jones, M. N. (2010). Evaluating the random representation assumption of lexical semantics in cognitive models. *Psychonomic Bulletin & Review*, *17*, 662–672.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115.

Jones, M. N., Hills, T., & Todd, P. M. (2015). Hidden processes in structural representations: A reply to Abbott, Austerweil, and Griffiths (2015). *Psychological Review*, *122*(3), 570–574.

Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences*, *14*(5), 223–232.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, *105*(31), 10687–10692.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, *68*(8), 1665–1692.

Koen, J. D., & Rugg, M. D. (2019). Neural dedifferentiation in the aging brain. *Trends in Cognitive Sciences*, *23*(7), 547–559.

Kovács, L., Bóta, A., Hajdu, L., & Krész, M. (2021). Networks in the mind–what communities reveal about the structure of the lexicon. *Open Linguistics*, *7*(1), 181–199.

Kumar, A. A. (2021). *Modeling semantic structure and spreading activation in retrieval tasks* (PhD thesis). Washington University in St. Louis.

Larkey, L. B., & Markman, A. B. (2005). Processes of similarity judgment. *Cognitive Science*, *29*(6), 1061–1076.

Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A., Weinberger, D. R., & Mattay, V. S. (2012). Normal age-related brain morphometric changes: Nonuniformity across cortical thickness, surface area and gray matter volume? *Neurobiology of Aging*, *33*(3), 617.e1–617.e9.

Li, Y., Engelthaler, T., Siew, C. S., & Hills, T. (2019). The Macroscope: A tool for examining the historical structure of language. *Behavior Research Methods*, *51*(4), 1864–1877.

Li, Y., & Siew, C. S. (2022). Diachronic semantic change in language is constrained by how people use and learn language. *Memory & Cognition*, *50*(6), 1284–1298.

Loesche, P., Wiley, J., & Hasselhorn, M. (2015). How knowing the rules affects solving the raven advanced progressive matrices test. *Intelligence*, *48*, 58–75.

Lord, L.-D., Stevner, A. B., Deco, G., & Kringelbach, M. L. (2017). Understanding principles of integration and segregation using whole-brain computational connectomics: Implications for neuropsychiatric disorders. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *375*(2096), 20160283.

Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and cognitive functioning across the life span. *Psychological Science in the Public Interest*, *21*(1), 6–41.

Luef, E. M. (2022). Growth algorithms in the phonological networks of second language learners: A replication of siew and vitevitch (2020a). *Journal of Experimental Psychology: General*, *151*(12), e26.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, *130*(2), 199–207.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, *22*(4), 796–810.

May, R. M. (2013). Networks and webs in ecosystems and financial systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1987), 20120376.

McCarley, J. S., Yamani, Y., Kramer, A. F., & Mounts, J. R. (2012). Age, clutter, and competitive selection. *Psychology and Aging*, *27*(3), 616–626.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407.

Menczer, F., Fortunato, S., & Davis, C. A. (2020). *A first course in network science*. Cambridge University Press.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, *117*(3), 363–386.

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), 758–799.

Oberauer, K., & Lin, H.-Y. (2024). An interference model for visual and verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *50*(6), 858–888.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*, 173–196.

Peper, J. S., Brouwer, R. M., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (2007). Genetic influences on human brain structure: A review of brain imaging studies in twins. *Human Brain Mapping*, *28*(6), 464–473.

Pietschnig, J., Gerdesmann, D., Zeiler, M., & Voracek, M. (2022). Of differing methods, disputed estimates and discordant interpretations: The meta-analytical multiverse of brain volume and IQ associations. *Royal Society Open Science*, *9*(5), 211621.

Qiu, X., FM Oliveira, D., Sahami Shirazi, A., Flammini, A., & Menczer, F. (2017). Limited individual attention and online virality of low-quality information. *Nature Human Behaviour*, *1*(7), 1–7.

Ramscar, M., Hendrix, P., Love, B., & Baayen, R. H. (2013). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *The Mental Lexicon*, *8*(3), 450–481.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, *6*(1), 5–42.

Ramscar, M., Sun, C. C., Hendrix, P., & Baayen, H. (2017). The mismeasurement of mind: Life-span changes in paired-associate-learning scores reflect the "cost" of learning, not cognitive decline. *Psychological Science*, *28*(8), 1171–1179.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*(3), 151–160.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Prokasy (Ed.), *Classical conditioning II: Current research and theory* (pp. 64–69). Appleton-Century-Crofts.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *Journal of Neuroscience*, *23*(8), 3295–3301.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Current Directions in Psychological Science*, *17*(3), 177–182.

Riedel, L., Heuvel, M. P. van den, & Markett, S. (2022). Trajectory of rich club properties in structural brain networks. *Human Brain Mapping*, *43*(14), 4239–4253.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*(2), 358–367.

Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience*, *35*(7), 1190–1200.

Rotaru, A. S., Vigliocco, G., & Frank, S. L. (2018). Modeling the structure and dynamics of semantic processing. *Cognitive Science*, *42*(8), 2890–2917.

Salthouse, T. A. (1992). *Mechanisms of age-cognition relations in adulthood.* Lawrence Erlbaum Associates, Inc.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144.

Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, *34*(1), 17–24.

Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of Clinical and

Experimental Neuropsychology, *25*(5), 625–633.

Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B., & Gouvier, W. D. (2010). The relationships of working memory, secondary memory, and general fluid intelligence: Working memory is special. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 813–820.

Siew, C. S. (2019). spreadr: An R package to simulate spreading activation in a network. *Behavior Research Methods*, *51*(2), 910–929.

Siew, C. S., & Vitevitch, M. S. (2020). An investigation of network growth principles in the phonological language network. *Journal of Experimental Psychology: General*, *149*(12), 2376–2394.

Siew, C. S., Wulff, D. U., Beckage, N., & Kenett, Y. N. (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. *Complexity*, *2019*, 2108423.

Simon, H. A. (1977). The organization of complex systems. In *Models of discovery: And other topics in the methods of science* (pp. 245–261). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-9521-1_14

Soto, F. A., Vogel, E. H., Uribe-Bahamonde, Y. E., & Perez, O. D. (2023). Why is the Rescorla-Wagner model so influential? *Neurobiology of Learning and Memory*, *204*, 107794.

Spreng, R. N., & Turner, G. R. (2019). The shifting architecture of cognition and brain function in older adulthood. *Perspectives on Psychological Science*, *14*(4), 523–542.

Spreng, R. N., & Turner, G. R. (2021). From exploration to exploitation: A shifting mental mode in late life development. *Trends in Cognitive Sciences*, *25*(12), 1058–1071.

Stella, M. (2020). Multiplex networks quantify robustness of the mental lexicon to catastrophic concept failures, aphasic degradation and ageing. *Physica A: Statistical Mechanics and Its Applications*, *554*, 124382.

Stella, M., Beckage, N., & Brede, M. (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports*, *7*(1), 1–10.

Stern, Y., Gurland, B., Tatemichi, T. K., Tang, M. X., Wilder, D., & Mayeux, R. (1994). Influence of education and occupation on the incidence of Alzheimer's disease. *Jama*, *271*(13), 1004–1010.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41–78.

Stoltzfus, E. R., Hasher, L., & Zacks, R. T. (1996). Working memory and aging: Current status of the inhibitory view. *Working Memory and Human Cognition*, 66–88.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268–276.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*(2), 135–170.

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

Trimmer, P. C., McNamara, J. M., Houston, A. I., & Marshall, J. A. (2012). Does natural selection favour the Rescorla–Wagner rule? *Journal of Theoretical Biology*, *302*, 39–52.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*(1), 138–146.

Tucker-Drob, E. M., De la Fuente, J., Köhncke, Y., Brandmaier, A. M., Nyberg, L., & Lindenberger, U. (2022). A strong dependency between changes in fluid and crystallized abilities in human cognitive aging. *Science Advances*, *8*(5), eabj2422.

Turnbull, L., Hütt, M.-T., Ioannides, A. A., Kininmonth, S., Poeppl, R., Tockner, K., et al.others. (2018). Connectivity and complex systems: Learning from a multi-disciplinary perspective. *Applied Network Science*, *3*, 1–49.

Unsworth, N. (2019). Individual differences in long-term memory. *Psychological Bulletin*, *145*(1), 79–139.

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, *18*(2), 332–339.

Vidal-Pineiro, D., Wang, Y., Krogsrud, S. K., Amlien, I. K., Baaré, W. F., Bartres-Faz, D., et al.others. (2021). Individual variations in "brain age" relate to early-life factors more than to longitudinal brain change. *Elife*, *10*, e69995.

Vitevitch, M. S., & Mullin, G. J. (2021). What do cognitive networks do? Simulations of spoken word recognition using the cognitive network science approach. *Brain Sciences*, *11*(12), 1628.

Weeks, J. C., Grady, C. L., Hasher, L., & Buchsbaum, B. R. (2020). Holding on to the past: Older adults show lingering neural activation of no-longer-relevant items in working memory. *Journal of Cognitive Neuroscience*, *32*(10), 1946–1962.

Wulff, D. U., De Deyne, S., Aeschbach, S., & Mata, R. (2022). Using network science to understand the aging lexicon: Linking individuals' experience, semantic networks, and cognitive performance. *Topics in Cognitive Science*, *14*(1), 93–110.

Wulff, D. U., De Deyne, S., Jones, M. N., & Mata, R. (2019). New perspectives on the aging lexicon. *Trends in Cognitive Sciences*, *23*(8), 686–698.

Wulff, D. U., Hills, T., & Mata, R. (2022). Structural differences in the semantic networks of younger and older adults. *Scientific Reports*, *12*(21459), 1–13.

Yau, J. O.-Y., & McNally, G. P. (2023). The rescorla-wagner model, prediction error, and fear learning. *Neurobiology of Learning and Memory*, *203*, 107799.

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, *22*(1), e12735.

Zhu, J.-Q., Sundh, J., Spicer, J., Chater, N., & Sanborn, A. N. (2024). The autocorrelated bayesian sampler: A rational process for probability judgments, estimates, confidence intervals, choices, confidence judgments, and response times. *Psychological Review*, *131*(2), 456–493.

Zortea, M., Menegola, B., Villavicencio, A., & Salles, J. F. de. (2014). Graph analysis of semantic word association among children, adults, and the elderly. *Psicologia: Reflexao* *e Critica*, *27*, 90–99.

Table 1

Statistics for the environments and growing representations.

Type	Measures	Environment	Epoch 1	Epoch 2	Epoch 3	Epoch 4
ER random	Nodes	499.84	305.15	415.16	459.33	478.76
	Edges	1984.16	234.82	449.25	941.92	1832.35
	Strength	8.003	0.010	0.020	0.029	0.039
	Degree	7.94	0.94	1.80	3.77	7.33
	ASPL	3.239	0.056	0.091	0.033	0.018
	C	0.016	0.002	0.004	0.019	0.043
	M	0.282	0.952	0.766	0.647	0.583
Scale-free	Nodes	442.23	172.12	248.13	295.28	327.32
	Edges	1390.60	219.51	755.67	1570.71	2430.35
	Strength	9.042	0.009	0.016	0.022	0.028
	Degree	6.29	0.99	3.42	7.10	10.99
	ASPL	3.141	0.028	0.016	0.012	0.011
	C	0.382	0.158	0.182	0.201	0.214
	M	0.111	0.597	0.535	0.500	0.475
Small-world	Nodes	499.81	305.43	415.35	459.63	478.88
	Edges	1941.61	234.30	447.05	936.21	1803.15
	Strength	8.003	0.010	0.020	0.029	0.039
	Degree	7.77	0.94	1.79	3.75	7.22
	ASPL	3.551	0.050	0.099	0.034	0.019
	C	0.049	0.006	0.012	0.028	0.051
	M	0.653	0.957	0.800	0.707	0.670
Scale-free small-world	Nodes	445.21	176.64	253.54	301.05	333.10
	Edges	1193.40	199.44	666.05	1414.21	2225.90
	Strength	8.981	0.009	0.017	0.024	0.030
	Degree	5.36	0.90	2.99	6.35	9.99
	ASPL	3.612	0.038	0.019	0.014	0.012
	C	0.379	0.165	0.190	0.208	0.223
	M	0.611	0.744	0.717	0.706	0.700

Note: Measures averaged over 1000 environments and cognitive representations learned over four epochs of 250 learning events each. Nodes indicate non-isolates. Strength is sum of edge weights. Degree is number of associations positive. ASPL is average shortest path length. C is mean local clustering coefficient. M is modularity.