1 A network enrichment account of cognitive aging

2 Thomas Hills[1]

3 [1] University of Warwick

4 Author Note

7 Correspondence concerning this article should be addressed to Thomas Hills, Gibbet

8 Hill Road, Coventry, CV4 7AL, UK. E-mail: t.t.hills@warwick.ac.uk

9   Abstract

10 Late-life cognitive development is associated with a decline in fluid intelligence which

11 occurs alongside a corresponding increase in crystallized intelligence. Theory often treats

12 these accounts independently, seeing the former as a consequence of biological aging and

13 the latter as a consequence of learning. What has not been fully explored is that lifelong

14 learning may explain both accounts. This article describes a formal enrichment account

15 that shows how a lifetime of experience encodes a cognitive representation that, when used

16 to guide behavior, produces numerous quantitative effects commonly described for

17 cognitive aging: with age, similarity judgments between concepts fall, free associations

18 become less predictable (higher entropy), and free association networks produced by

19 aggregating output across associates show patterns of declining connectivity, indicated by

20 falling network degree, rising average shortest path lengths, and a reduction in local

21 clustering. The enrichment account demonstrates how these behavioral outcomes arise

22 when a general prediction error model (Rescorla-Wagner) is used to train a cognitive

23 representation through repeated experience with a structured environment and the

24 cognitive representation is then sampled from to produce behavior (e.g., free associates or

25 similarity judgments). This diffusion of activity and its associated consequences is a

26 general property of network enrichment which also speaks directly to cognitive aging.

27 Moreover, as measures of co-activation, rising entropy and falling similarity judgments

28 provide mechanisms for cognitive slowing, explaining declining fluid intelligence. These

29 results help better inform biological mechanisms of aging and its associated pathologies by

30 better calibrating our predictions for accounts based on atrophy and enrichment.

31   *Keywords:* cognitive aging, Rescorla Wagner, spreading activation, network science,

32 free associations, fluid intelligence, crystallized intelligence, cognitive slowing

A network enrichment account of cognitive aging

Cognitive aging across the adult lifespan 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 increase (Brysbaert, Stevens, Mandera, & Keuleers, 2016; Park & Reuter-Lorenz, 2009; Salthouse, 2004). This distinction 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*, is a classic division of intelligence (Cattell, 1987). This division also characteristically distinguishes the old from the young. What explains these age-related changes? Might they be explained by a single process? The evidence I provide below suggests that the answer is yes: the outcome of enriching one's cognitive representation through a lifetime of experiencing associations in the world produces an average decline in activation between any two concepts chosen at random. Before demonstrating how this effect arises, it is first useful to highlight the focal alternative accounts of aging this work addresses and to describe the behavioral effects we hope to explain.

Degradation or enrichment

Cognitive aging is a rich process with evidence supporting a wide range of phenomenology and theoretical explanations (Cao et al., 2014; Grady, Springer, Hongwanishkul, McIntosh, & Winocur, 2006; Koen & Rugg, 2019; Mata, Schooler, & Rieskamp, 2007; e.g., Reuter-Lorenz & Cappell, 2008; Spreng & Turner, 2021; Troyer, Moscovitch, & Winocur, 1997). Much of this work is often interpreted as supporting either degradation-based accounts (e.g., involving declining frontal lobe or executive function and neural atrophy) or enrichment-based accounts (e.g., involving adaptive consequences of

lifelong learning), and sometimes both. This dichotomy between things falling apart or things coming together is well characterized by the distinction between fluid and crystallized intelligence.

Declines in fluid intelligence are often seen as independent of rising crystallized intelligence. One prominant explanation for declines in fluid intelligence 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 (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 mechanisms underpinning cognitive performance. Salthouse (1992) describes some potential mechanisms: "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). Cortical thickness also declines alongside increases in cerebrospinal fluid space (Lemaitre et al., 2012). These findings are often associated with atrophy and fit with common intuitions for biological aging—as biological things fall apart, so do the cognitive counterparts which they underpin. Moreover, neuropathology—associated with posthumously verified evidence of Alzheimer's and other neurodegenerative and cerebrovascular disease—accounts for up to 40% of the variation in late-life cognitive decline (Boyle et al., 2021). Though this leaves substantial variance in cognitive-decline unexplained, it nonetheless identifies an important correspondence between biology and cognition, especially in unhealthy aging.

Alternatively, and far more rarely, some accounts of cognitive aging have proposed pathways suggesting an interdependence between the positive consequences of learning (crystallized intelligence) and their negative consequences for fluid intelligence. For example, Buchler and Reder (2007) showed using simulations that if they assumed that the

number of relations between concepts increased over the lifespan this would lead to more diffuse activation between concepts, analogous to a contextual fan effect. This reproduced age-related changes in recognition and familiarity. Using a different approach, Ramscar, Hendrix, Shaoul, Milin, and Baayen (2014) demonstrated how associations would change through repeated learning. They based their work on Desrosiers and Ivison (1986) study of paired-associate learning in older adults, which found that older adults perform most poorly on stimuli that are least consistent with their prior experience. Ramscar, Sun, Hendrix, and Baayen (2017) showed that the difficulty of learning unrelated word pairs is entirely predictable from the co-occurrence frequency of those pairs in past experience. Training a Rescorla-Wagner model on typical patterns of word co-occurrences, unrelated word pairs become negatively associated over time, impairing the future learning of their association.

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 executive function at the time of encoding or retrieval (Amer, Wynn, & Hasher, 2022). According to this account, processing deficits and the inability to filter out past experience can lead older adults to attend to too much information. This in turn creates processing difficulties when that information turns out to be irrelevant. Though there is no existing formal model of clutter development, the enrichment account provided below offers one.

The different accounts describe above can be broadly categorized as degradation versus enrichment accounts, and the evidence for both is compelling. However, to what extent we subscribe to one explanation over another should largely depend on how they work and what they explain. To my knowledge, the 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) have developed a compelling model of degradation for hyper-priming in

Alzheimer's patients, but similar models for healthy aging are few and far between. In addition, formal degradation models would need to explain how degradation alters fluid intelligence but not crystallized intelligence.

On the other hand, enrichment explanations either assume more connectivity (e.g., Buchler and Reder (2007)) or evaluate how differential experience impairs or enhances pair-wise associations (e.g., Ramscar et al. (2014)). As called for elsewhere (Wulff, De Deyne, Jones, & Mata, 2019), what is lacking in both cases is a full model of representational development and behavior across the adult lifespan that takes into account our understanding of the aging lexicon. This would represent a computational prediction for what we should expect from lifelong enrichment and what remains to be explained by degradation. To benchmark such a model, we can use a number of empirical observations made over the last decade.

The aging lexicon, entropy, and similarity

Several efforts to chart the mental lexicon across the lifespan have identified reproducible trends. Dubossarsky, De Deyne, and Hills (2017) asked over 8000 people, ranging in age from roughly 10 to 70, to provide three free associates to each of 420 words. With approximately 1000 people in each age group, data was aggregated within age-groups to produce networks among the 420 words with edges representing a weighted function of common associates. After removing edges below a threshold, older networks were found to have a lower average degree (number of associations per word) higher average shortest path length (greater distance between associates), and lower average clustering coefficient (the proportion of a concept's neighbors that are themselves connected). Evidence also identified a rising entropy for associations, with associates becoming less predictable across the lifespan. Both Zortea, Menegola, Villavicencio, and Salles (2014) and Wulff, De Deyne, Aeschbach, and Mata (2022), with varying numbers of participants and cues and barring differences in network construction methods, found similar patterns.

Analyses of memory search in older and younger adults also find consistent patterns of change in the aging mental lexicon. Using a semantic fluency task ("name all the animals you can think of"), Wulff, Hills, and Mata (2022) constructed lexical networks by connecting words that appeared nearby in the lists that older and younger adults produced. This revealed that older adults' lexicons were less well-connected, similar to the patterns for free associations described above. Hills, Mata, Wilke, and Samanez-Larkin (2013) modeled semantic fluency data from adults across the lifespan and found that, compared with younger adults, older adults produced strings of less similar words. Finally, Cosgrove, Kenett, Beaty, and Diaz (2021) used percolation analysis to investigate the resilience of older adults' mental lexicons by artificially removing connections between words and found that older adults' lexicons were less resilient to decay. These findings are all suggestive of more sparse connectivity in the outputs retrieved from the mental lexicon.

Slower judgments of similarity between words are equally consistent with the above patterns. Wulff, Hills, et al. (2022) asked younger and older adults to judge the similarity of 77 different animals over a period of several weeks using a tablet participants took to their homes. Rating the similarity of pairs of animals on a scale from 1 to 20, Wulff, Hills, et al. (2022) found that older adults rated animals as less similar to one another than did young adults.

The key take-away with respect to what follows is that older adults produce less predictable associations (higher entropy), lower similarity judgments, and show evidence of sparsening free association networks as measured by degree, average shortest path length, and clustering coefficient. These results are intuitively consistent with representational degradation: one may imagine that a sparseness in output reflects a sparseness in the underlying representation, which is in turn caused by degradation in the neural architecture that underpins it. However, without understanding what we might expect from lifelong learning, efforts to explain cognitive aging as the result of degradation may attempt to bridge an explanatory gap that does not exist. Moreover, degradation accounts

165 may even get the causation backwards. Which is to say, if life experience gives rise to some

166 of the primary markers of age-related cognitive decline, then so-called atrophy in

167 physiological structure may not be the cause of age-related cognitive decline. Rather, it

168 may be the consequence of life-long learning. As demonstrated below, by extending

169 standard learning and retrieval models across the lifespan, we can predict all of the above

170 effects as a consequence of enrichment, without the need for assuming any additional

171 processes associated with biological aging or degradation.

The enrichment account

173 The enrichment account envisions behavior as the outcome of learning relationships

174 from the environment to develop a cognitive representation, and then using this

175 representation to generate behavior. This follows calls from previous researchers to model

176 not only the representation but the processes that access that representation to generate

177 behavior (Castro & Siew, 2020; Estes, 1975; Hills & Kenett, 2022; Johns, Jamieson, &

178 Jones, 2023; Jones, Hills, & Todd, 2015). Critically, this account sees behavior as the

179 outcome of process and representation, not as a direct copy of the representation.

180 Formally, the enrichment account involves modeling three separate components:

181 1. *Environment*: The environment presents the set of possible associations that could be

182 learned.

183 2. *Representation*: Learning associations from the environment generates a cognitive

184 representation. This continues to develop across the lifespan.

185 3. *Behavior*: Behavior is recovery of information from the cognitive representation

186 appropriate to the environmental context. This generates free-associations, memory

187 search, similarity judgments, and so on.

188 These stages are presented in Figure 1 and each are explained in detail below. The R

189 code to reproduce the environments, learning of cognitive representations, behavior, and all

figures in this manuscript are available at https://github.com/thomasthills/enrichment.
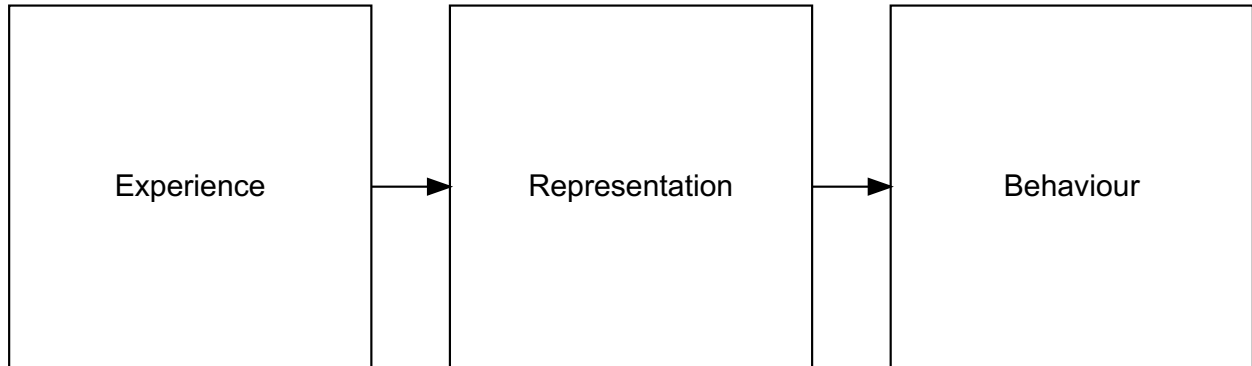

Figure 1. The process of translating environmental experiences into behaviour. Arrows represent processes that translate one domain into another. Learning translates experience into a cognitive representation. Additional cognitive processes then act on the representation to generate behavior.

Environment

The environment is represented here as network of concepts (e.g., words), with the edges (links) between concepts indicating associations that can be learned. This follows the basic idea that things in the world predict one another, and learning is here taken to address one association at a time. The qualitative pattern of behavioral results described below are not dependent on the specific network structure. To demonstrate this, I use a general framework for network construction. This is a variation of a fitness-based network model (Menczer, Fortunato, & Davis, 2020) using rank-based sampling. Each concept is assigned a rank, r, from 1 to 1000. Then pairs of concepts are chosen from the lexicon each with probability $p \propto r^{-a}$ and an edge is created between them. When $a = 1$, this process produces a scale-free network, which is characterized by a long-tail in which most concepts have very few associations and a few have very many. This structure is consistent with the ubiquity of scaling laws in the cognitive sciences and the natural world (Johns & Jones, 2010; Kello et al., 2010). When $a = 0$ this process approximates a weighted Erdös-Renyi

205 random graph, in which all nodes are connected with equal probability. Each network is
206 produced starting with 500 concepts. Then 2000 pairs of concepts are sampled (with
207 replacement) and weighted edges are formed between them, with the weight corresponding
208 to the number of times the pair is chosen. The edges are weighted and non-directional
209 (undirected). Figure 2 presents two example networks and their degree distributions for the
210 different values of a. The exact number of nodes and edges is not critical to this
211 demonstration and can be explored in the online code.

Representation

213 Cognitive representations are built from experience with the environment. Here, I use
214 the prediction error framework set out in the Rescorla-Wagner model (Rescorla & Wagner,
215 1972) to encode properties of the environment into a cognitive representation. The choice
216 of the Rescorla-Wagner model follows the strong evidence for learning as a process of
217 minimizing prediction error, which is a fundamental assumption among models of
218 reinforcement learning (Dayan & Abbott, 2005; Hoppe, Hendriks, Ramscar, & Rij, 2022;
219 McClelland & Rumelhart, 1981; Sutton & Barto, 2018). The Rescorla-Wagner model
220 captures this phenomenology—including associative learning, blocking, inhibition, and
221 extinction—and it is a model on which many subsequent models have been based (e.g.,
222 Sutton & Barto, 1981; Trimmer, McNamara, Houston, & Marshall, 2012). Though it is not
223 without limitations (Miller, Barnet, & Grahame, 1995; e.g., Yau & McNally, 2023), I vary
224 some of these limitations in the Supplementary Material. I therefore use the model to
225 capture the generic prediction-error process inherent in its design and in recognition of its
226 extreme predictive utility across a wide variety of research domains (Miller et al., 1995;
227 Roesch, Esber, Li, Daw, & Schoenbaum, 2012; e.g., Soto, Vogel, Uribe-Bahamonde, &
228 Perez, 2023).

229 Formally, the Rescorla-Wagner model minimizes the prediction error between the
230 values of an observed outcome, j, and a cue predictive of that outcome, i. The value

associated with 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, $(\lambda_j - V_i)$, 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 (some outcomes are learned about faster 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. This follows exactly the formalization set out in prior work (Ramscar et al., 2017; Rescorla & Wagner, 1972).

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 is represented in the environment). Of the two concepts associated with the relationship, one is randomly assigned as the cue and the other as the outcome. The representation is then updated according to the Rescorla-Wagner model with $\alpha = 1$ and $\lambda = 1$. To demonstrate that the qualitative results are not dependent on the precise learning values, β is varied from .1 to .5, which varies over the lower end of the range for the product of α and β common to this model. Indeed, the results are the same for much smaller values of β.

To simulate development, learning occurs in 4 epochs each with 200 learning trials. The precise number of epochs or learning trials is also unrelated to the results, as can be

252 explored in the online code. These values are chosen here largely for computational

253 simplicity. Figure 3 provides examples of two learning representations over the four epochs

254 for the different network types ($a = 0$ and $a = 1$) and

Results: Behavior

256   The stylized facts associated with cognitive aging are rising entropy, a reduction in

257 pairwise similarity judgments, and changes in the structure of the free association network.

258 Each of these is recovered from the representation as described below. In each case,

259 environments are constructed independently 1000 times and then learning occurs over 4

260 epochs of 200 trials each.

261   **Rising entropy.**   Rising entropy refers to the reduction in the predictability of free

262 association targets as individuals age (Dubossarsky et al., 2017). We can measure this

263 using *Shannon's information entropy*. This measures the surprisingness of associates given

264 the cue. If a cue has only a few strong associations, any given associate will be less

265 surprising than if it has many equally weighted associations. Because the output of

266 Rescorla-Wagner learning is a weighted edge, we can compute this for every cue in the

267 network representation as follows:

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

268   Here, p_i is the proportion of the weight along edge i for all k edges for that cue. That

269 is, $p_i = \frac{w_i}{\sum_k w_k}$. The entropy for the network reported below is the mean across all cues.

270 This is taken here to represent the long-term entropy of persons producing repeated

271 associations for each cue.

272   The result of this computation is shown in Figure 4. For various network

273 constructions and learning rates, the results consistently show that entropy rises with

274 increased learning, following that observed in Dubossarsky et al. (2017).

275 **Similarity.** To simulate similarity judgments, we can create a measure of

276 co-activation between cues. To do this, we allow spreading activation to leave one node and

277 measure activation at the other node, $A_{j \to k}$. This allows us to measure the extent to which

278 one word co-activates the other. Doing this for both cues, we take similarity as the

279 summed co-activation.

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

280 We measure this similarity for a random selection of 20 concept pairs in the

281 representation, all of which are learned during the first epoch of learning. This ensures that

282 individuals at each age (epoch) have experience with the concepts.

283 The result of this computation is shown in Figure 4. For various network

284 constructions and learning rates, the results consistently show that similarity judgments

285 fall with increased learning, following that found in Wulff et al. (2019).

286 `## Warning in xy.coords(x, y, xlabel, ylabel, log): 52 y values <= 0 omitted from`

287 `## logarithmic plot`

288 **Free association networks**

289 As noted above, Dubossarsky et al. (2017) and others found that free association

290 networks collected across the adult lifespan showed patterns of decreasing network degree,

291 increasing average shortest path length, and decreasing clustering coefficient. We simulate

292 this on the cognitive representations across the four epochs as follows. For each cognitive

293 representation, we simulate 10 participants who retrieve 3 associates from each of 30 cues.

294 The three associates are sampled without replacement for each participant and each cue is

295 sampled in proportion to the associative strength encoded in the cognitive representation,

296 i.e., the output from the Rescorla-Wagner training epochs. Negative association strengths
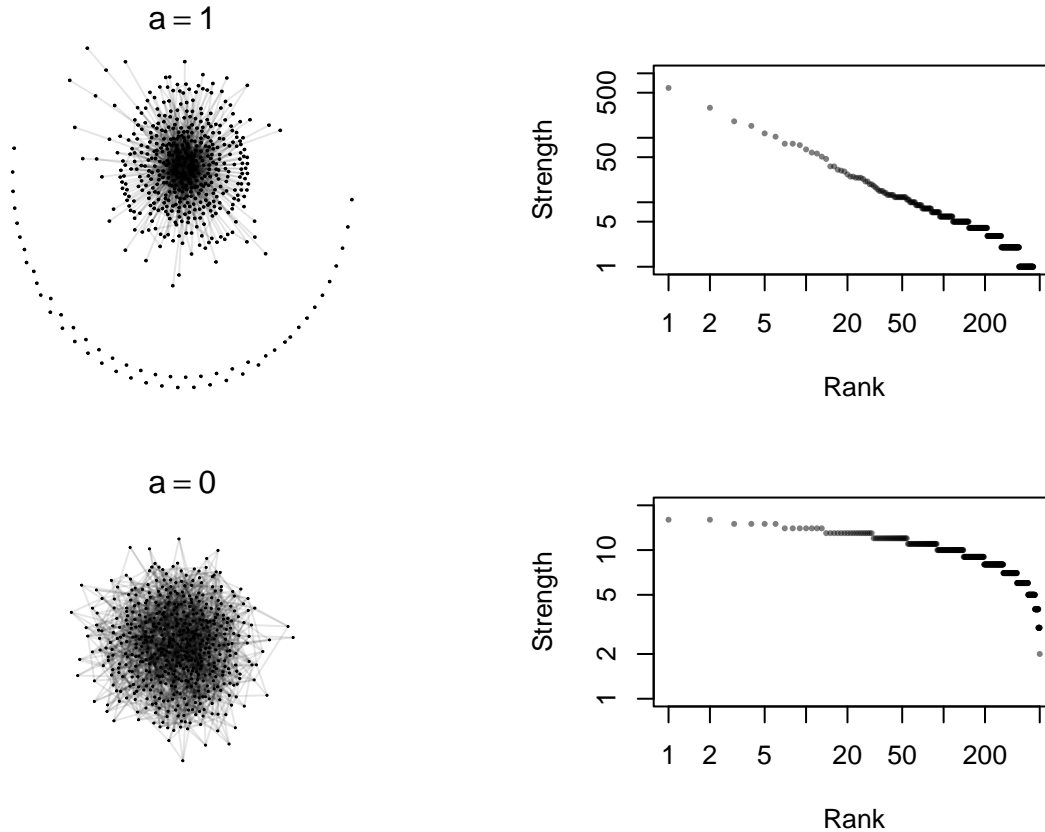
Figure 2. The structure of the environment for two network types: a scale-free network ($a = 1$) and a weighted Erdös-Renyi random graph ($a = 0$). Each network has 500 concepts (or nodes) and the weighted edges between them are the result of repeatedly sampling 2000 pairs of nodes and adding 1 to the edge weight between the pairs. The degree distributions are shown to help communicate the difference in the underlying structure.

297 are set to 0. This produces a cue by associate matrix, with each cell indicating the number

298 of times each associate was produced in response to each cue. From this matrix, cue-by-cue

299 similarities are computed using cosine similarity of their corresponding associate vectors.

300 To transform these into unweighted and undirected networks, the median cue-by-cue

301 similarity is computed across all epochs and, for each age, cue similarities below this

302 median value are set to 0, and all other values are set to 1. This matrix is then

303 transformed into a network on which degree (the number of connections), average shortest

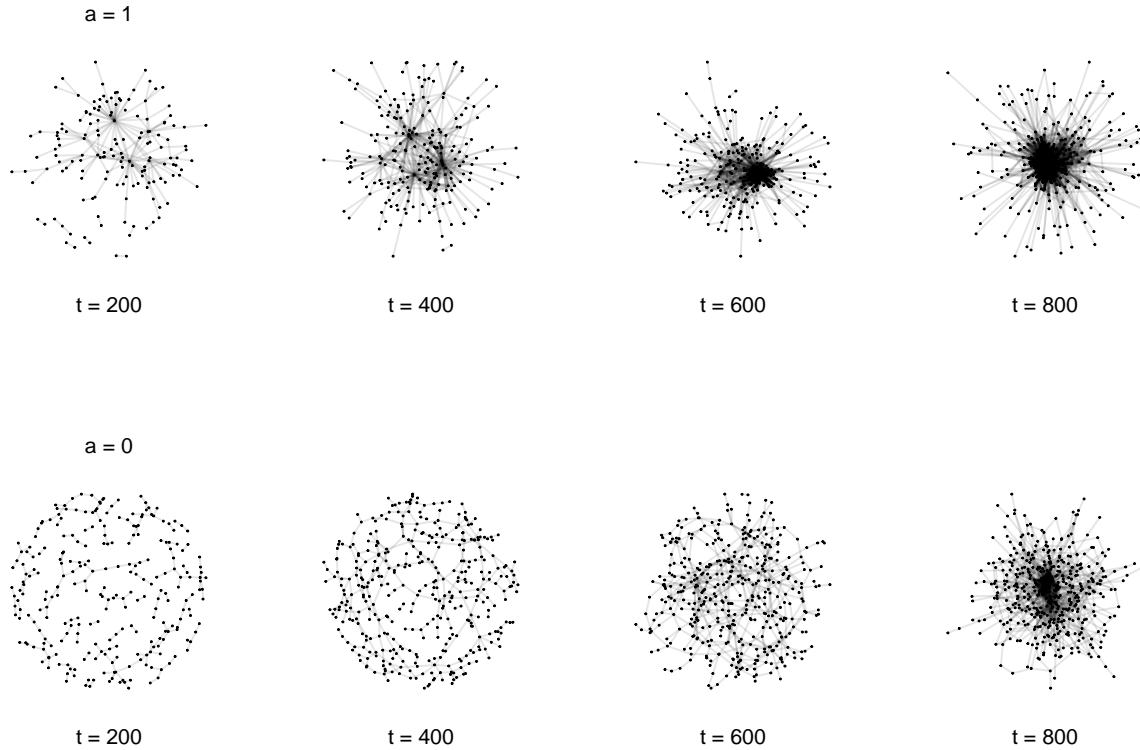304 path length (the shortest number of edges between cues), and clustering coefficient (the

Figure 3. Examples of the growing mental lexicon resulting from training a Rescorla-Wagner model on the network types shown in Figure 2. Training occurs in 200 event epochs, with edges from the environment sampled in proportion to their weight. Nodes represent individual concepts and edges represent learned associations.

proportion of a node's neighbors that are themselves connected) are computed. This entire process is repeated for 1000 different initial starting environments. The results are shown in Figure 5. The results follow the pattern found by Dubossarsky et al. (2017) and others.

Limitations and robustness

The enrichment account is a general framework that demonstrates how encoding an age-enriched representation that is then used to generate behavior is sufficient to generate several key empirical trends in the aging literature. This is subject to limitations, some of which can be addressed here, and other motivate additional research. These are listed in no particular order below.
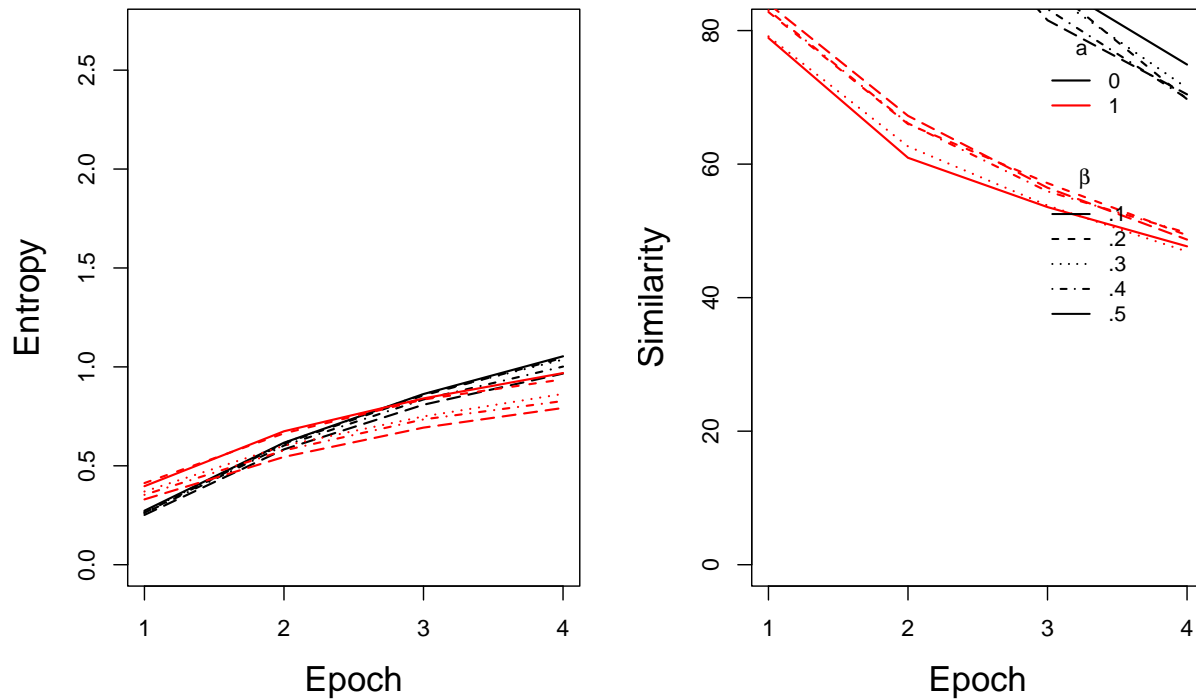
Figure 4. Entropy and similarity measures computed from the cognitive representation at different epochs of development.

- *Number of nodes or edges* The estimate of the number of concepts people experience varies widely. A factor of 10 estimate is generally around 100,000 (Brysbaert et al., 2016), though it could be much more or far less depending on one's environment (Hart & Risley, 1995). As many language researchers note, the number could be infinite, since new words are unlimited. In addition, the number of relationships between concepts is in principle unknowable and limited only by the network size. In practice, humans report only a subset of potential associations, ???. For the purposes of demonstration, I have chosen a fairly small vocabulary, which makes this work computationally tractable and easy for researchers to test with the online code. That said, the supplementary material shows that large changes in network size and density (number of associations over the number of possible associations) generate
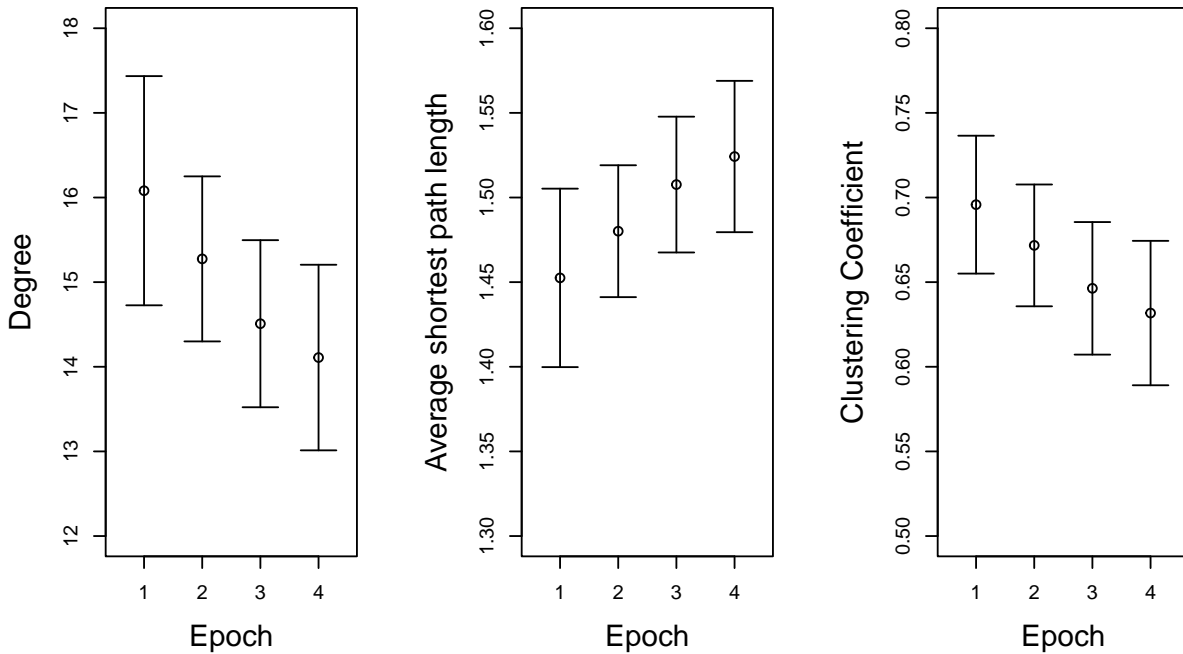
Figure 5. The degree, average shortest path length, and clustering coefficient for networks of associations across the lifespan. Simulations were repeated for learning cognitive representations in 1000 different environments ($a = 1$) with four training epochs each of 200 learning events each($beta = .1$). Representations were then each sampled from by 10 simulated participants who each retrieved 3 associates for each of 30 cues with probability proportional to the associative strengths output from the Rescorla-Wagner training epochs. The cosine similarity of cues was then computed from the cue by association matrix to produce associative networks for each epoch. These networks were then thresholded separately for each learning environment by the median similarity value across all ages. Error bars indicate one standard deviation. Results are consistent with Dubossarsky et al., 2017.

325   the pattern presented here.

326   • *Existence of context cue in Rescorla Wagner* The presence of the always present

327   'context' cue has been criticized as creating excitatory activation among cues that

328   never occur with downstream targets. As shown in the supplementary material,

329   removal of this cue does not influence the behavioral patterns described here.

330   • *Learning parameters* Learning parmaters are explored over a range of values in the

331   SM, and consistently demonstrate the behavior described here.

332   • *Number of learning trials* The appropriate number of learning trials is varied in the

333   SM and consistently shows the pattern produced here.

334   • *Herdan-Heaps law* The results of the enrichment account are also preservd if the size

335   of the environment expands with age, consistent with Herdan-Heaps' law.

336   • *Stages of development* This work captures adult development patterns and is not

337   meant to characterize child development. Though lexical development in children has

338   been addressed elsewhere (Ciaglia, Stella, & Kennington, 2023; Fourtassi, Bian, &

339   Frank, 2020; Jiménez & Hills, 2022), extending the enrichment account to children

340   remains an open question.

Discussion

342   Aging is marked by complex and multi-faceted phenomenology. The enrichment

343   account provided here demonstrates that a subset of this phenomenology is predicted by

344   cognitive enrichment. Specifically, the enrichment account demonstrates how learning

345   enriches the cognitive representation and then, using various processes, gives rise to

346   behavior commonly associated with cognitive aging. In particular, enrichment reproduces

347   rising entropy, falling similarity judgments, and a sparsening of the free association network

348   across the lifespan.

349   The enrichment account is not meant to be a complete explanation of aging. Rather,

350   like any model, it puts guardrails on our ignorance. It does this by showing how many of

351   the behaviors that might be intuitively interpreted as evidence of cognitive and neural

352   degradation are straightforward outcomes of learning. Where we might have suspected

353   degradation, enrichment may be at least equally likely. Indeed, Borge-Holthoefer et al.

354   (2011) used degradation *not* to demonstrate diffusion of activation and slowing, but rather

355   to demonstrate an increasing of activation between associates, producing hyper-priming, a

356   well-documented behavior in patients with Alzheimer's Disease. That said, the patterns of

357   biological and neural change outlined in the introduction are themselves data, and any full

358   account of cognitive aging requires that we understand how these interact with enrichment

359   to produce the behavioral variation that we see. With respect to parsimony, enrichment

360   may be sufficient to explain how fluid intelligence might change as a result of changes in

361   crystallized intelligence (as opposed to treating each independently), but our theories must

362   also be parsimonious with respect to observed changes in the underlying neural

363   architecture.

364   Importantly, the enrichment account is supported empirically. For example, the fan

365   effect demonstrates that learning many relationships with a target concept reduces the

366   speed of accessing any one of those relationships (Anderson & Reder, 1999). This is a

367   general and well-understood outcome of spreading activation or random walker models,

368   which lead to greater activation in targets when the pathways along which they must

369   channel activation are more limited (Abbott, Austerweil, & Griffiths, 2015; Siew, 2019). In

370   addition, one can see this effect in natural experiments: Ramscar et al. (2017) showed that

371   native speakers of a language were more impaired in paired-associative learning than

372   age-matched second language speakers (who had less experience with the language). This

373   suggests the effect is not about age, but about enrichment.

374   We also know that external clutter influences processing speed in older adults more

375   strongly than young adults (Amer et al., 2022; McCarley, Yamani, Kramer, & Mounts,

2012). This is consistent with higher conceptual entropy creating a more level-playing field for competition. If there is greater associative competition in older adults, because the associative entropy is higher, then older adults will suffer more from clutter in internal and external processing. External cues gain their relevance via their encoded associations in memory (Easdale, Le Pelley, & Beesley, 2019). The attentional contribution to clutter on aging memory, as elegantly described by Amer et al. (2022), is therefore explained by the enrichment account: the activation strengths associated with extraneous and potentially irrelevant information has a greater relative competitive advantage in individuals with more enriched representations, making older individuals more susceptible to their effects.

One potential criticism of the enrichment account is that individuals with higher education and occupational attainment are less likely to experience late-life cognitive decline (Lövdén, Fratiglioni, Glymour, Lindenberger, & Tucker-Drob, 2020; Stern et al., 1994). This is known to be partially a consequence of compensation accorded by skills or strategic repertoires, what is called cognitive reserve (Scarmeas & Stern, 2003). Additionally, however, there is ample evidence that differences in cognitive skill 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 differences in processing. Early processing advantages may not be a result of experience at all. Brain volume is correlated with IQ (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). Nonetheless, if we believe that formal education (or occupational attainment) is an indicator of increased associative density (not just in quality, but quantity), then an ideal test might be to investigate twins who differ in educational enrichment are then evaluated on age-related cognitive decline. The Ramscar

et al. (2017) study on age-matched bilinguals and monolinguals is another approach. What kinds of education matter for enrichment is an open question. Is rewatching the 11 seasons of the 1970s TV series M*A*S*H (approximately 125 hours runtime) different from reading the complete works of Gabriel Garcia Marquez? If the enrichment is about the variety of associations, and not the quality of associations (if such quality could be objectively measured), then the difference may not matter. To be alive at all is to experience enrichment.

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.

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.

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.

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.

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.

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

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.

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.

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

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.

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.

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.

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.

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., 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.

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.

Jiménez, E., & Hills, T. (2022). Semantic maturation during the comprehension-expression
gap in late and typical talkers. *Child Development*, *93*(6), 1727–1743.

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.

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.

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

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.

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.

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.

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

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.

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.

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., & 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.

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

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.

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.

Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of Clinical and Experimental Neuropsychology*, *25*(5), 625–633.

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

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

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.

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

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.

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.

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

572 fear learning. *Neurobiology of Learning and Memory, 203*, 107799.

573 Zortea, M., Menegola, B., Villavicencio, A., & Salles, J. F. de. (2014). Graph analysis of

574 semantic word association among children, adults, and the elderly. *Psicologia: Reflexao*

575 *e Critica, 27*, 90–99.