

An enrichment account of cognitive aging

Thomas Hills¹

¹ University of Warwick

Author Note

This work was supported by the Alan Turing Institute and Royal Society Wolfson Research Merit Award WM160074.

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

Abstract

10 Late-life cognitive development is associated with a decline in fluid intelligence alongside a
11 corresponding increase in crystallized intelligence. Though age-related cognitive decline in
12 fluid intelligence is often associated with a common-cause account of biological aging, what
13 has not been formally explored is that a rise in crystallized intelligence might explain a the
14 decline in fluid intelligence. Here I describe a model of learning across the lifespan that
15 shows how standard reinforcement learning exposed to a lifetime of associative learning can
16 produce two effects associated with cognitive aging: higher entropy in associative responses
17 and a fall in similarity judgments. As measures of co-activation, these also provide a
18 mechanism for cognitive slowing. The enrichment account assumes that individuals learn a
19 cognitive representation through repeated experience with a structured environment. They
20 then sample that representation using spreading activation to produce associates and make
21 similarity judgements. Standard effects of cognitive aging are, by this account, the
22 consequence of an enriched cognitive representation.

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

An enrichment account of cognitive aging

Cognitive aging across the adult lifespan is characterized by two distinct and well-documented patterns: as individuals age, many measures of working memory, processing speed, and long-term memory show apparent performance decrements from approximately the age of 20, while at the same time, measures of vocabulary and other kinds of general knowledge increase (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 knowledge*, is a classic division of intelligence (Cattell, 1987). Several recent accounts of cognitive aging have argued for a relationship between these two phenomena (Amer, Wynn, & Hasher, 2022; Buchler & Reder, 2007; e.g., Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014). This raises the challenge: can we provide a formal explanation for how a rise in crystallized intelligence facilitates a decline in fluid intelligence?

It is useful to juxtapose this interactionist account against an often proposed alternative that treats these two phenomena independently. For example, the *common cause theory of age-related cognitive decline* 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) offers detailed biological explanations for how this might work: “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). While this account fits with our intuition for biological aging, it is useful to ask if this account is necessary. Would cognitive enrichment cause cognitive slowing on its own?

How enrichment could lead to slowing has seen several formal treatments. Buchler and Reder (2007) modeled age-related changes in word recognition after a contextual fan effect, in which life-experience increases concept relations and leads to more diffuse activation. Their model manipulated base-level activation and number of contextual relations (fan) and demonstrated that these changes could reproduce age-related changes in word recognition. Ramsar et al. (2014) took a different approach to explain age-related declines in paired-associate learning. Based on Desrosiers and Ivison (1986) observation that older adults perform more poorly on paired-associate learning than younger adults, Ramsar et al. (2014) suggested that “older adults’ changing performance reflects memory search demands, which escalate as experience grows” (p. 5) because older adults largely show impaired paired-associated learning only for unrelated terms. In subsequent work, Ramsar, Sun, Hendrix, and Baayen (2017) showed that the difficulty of learning unrelated word pairs is entirely predictable from the frequency of co-occurrence of those words. Training a Rescorla-Wagner model on typical patterns of word co-occurrences, unrelated word pairs become negatively associated over time. As Ramsar et al. (2017) state, “the discriminative processes that produce ‘associative’ learning teaches English speakers not only which words go together, but also which words do not go together. This process both increasingly differentiates meaningful and meaningless word pairs and makes meaningless pairs harder to learn” (p. 3). 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 (Amer et al., 2022).

Recent research adds additional nuance to these findings based on free associations, memory search, and similarity judgments. First, several efforts to chart the mental lexicon across the lifespan using free associates have revealed reproducible patterns. 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. Across the adult lifespan, older networks had lower degree (number of associations), higher average shortest path length, and higher entropy for associations (less predictable associations). Zortea, Menegola, Villavicencio, and Salles (2014) found a similar pattern with a smaller group of participants. With a still smaller group of participants ($n=8$) but far more cues ($n=3000$), Wulff, De Deyne, Aeschbach, and Mata (2022) found this pattern yet again.

Analyses of memory search in older and younger adults also find consistent patterns of change in the aging mental lexicons. Using semantic fluency data—e.g., “name all the animals you can think of”—Wulff et al. (2022) found that the mental lexicons of older adults appeared less well-connected. This used edges based on nearby co-occurrence in the string of productions, producing networks with lower average degree and higher average shortest path lengths. Hills, Mata, Wilke, and Samanez-Larkin (2013) modeled the fluency task using semantic space models and found that older adults produced less similar words when searching memory than younger adults. Finally, Cosgrove, Kenett, Beaty, and Diaz (2021) used percolation analysis to investigate the resilience of older adults’ mental lexicons, finding that older lexicons were less resilient to decay than younger networks.

Finally, older adults also judge animals to be less similar to one another than younger adults. Wulff, De Deyne, Jones, and Mata (2019) also asked younger and older adults to judge the similarity of 77 different animals. Rating the similarity of pairs of animals on a scale from 1 to 20, Wulff et al. (2019) found that older adults tended to rate animals as less similar than young adults.

In sum, older adults produce less predictable associations (higher entropy) and lower similarity judgments than younger adults. These results are intuitively consistent with a common cause account of age-related decline, and possibly of representational degradation. However, might these results be a natural consequence of enrichment based on life-experience? extending standard learning and retrieval models across the lifespan, we

can predict all of these effects as a consequence of enrichment, without the need for assuming any additional processes associated with biological aging or degradation.

The enrichment model

The enrichment account envisions behavior as the outcome of three stages:

1. *Environment*: The environment presents the set of possible experiences, i.e., associations.
2. *Representation*: Learning from the environment generates a cognitive representation. This continues to develop across the lifespan.
3. *Behavior*: Behavior is recovery of information from the cognitive representation appropriate to the environmental context. This generates free-associations, memory search, similarity judgments, and so on.

These stages are presented in Figure ?? and each is explained in detail below.

Environment

The environment presents the set of possible relationships an individual could learn. It is represented here as network (matrix) cues, with the edges between cues indicating learnable associations. The environment presented here is a variation of fitness-based network model using rank-based sampling. This is inspired by the ubiquity of scaling laws in the cognitive sciences and the natural world (Kello2010ga?). This assumes a power-law relation in the frequency of cues: Each cue is assigned a rank, r , from 1 to 1000. Then pairs of cues are chosen from the lexicon with probability $p \propto r^{-a}$ and an edge is created between them. Here a is set to .1 and 2000 edges are created. A scale-free network is not required to get the aging result shown below—for example, an Erdős-Renyi random graph produces the same qualitative pattern. In addition, the number of words available to

learn could increase across the lifespan, as proposed by Brysbaert, Stevens, Mandera, and Keuleers (2016) and following *Herdan-Heaps' law* (Petersen, Tenenbaum, Havlin, Stanley, & Perc, 2012; Serrano, Flammini, & Menczer, 2009). Again, the qualitative results are the same. Examples are provided in the Supplementary material.

Representation

Contemporary accounts characterize learning as 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 (Rescorla & Wagner, 1972) captures this phenomenology and is a model on which many subsequent models have been based (e.g., Sutton & Barto, 1981; Trimmer, McNamara, Houston, & Marshall, 2012). Moreover, it captures phenomenology like association learning, blocking, inhibition, and extinction.

The Rescorla-Wagner model formalizes learning as a process of minimizing prediction error between the values of an observed outcome, λ_j and a predictive cue, V_i , where j and i represent specific outcomes and cues, respectively. The prediction error is the difference ($\lambda_j - V_i$) and it is minimized following each learning event according to the following rule:

$$\Delta V_{C \rightarrow U} = \alpha_C \beta_U (\lambda_U - V_{C \rightarrow U})$$

α corresponds to cue salience (some cues are easier to learn about than others) and β 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

$$V_{C \rightarrow U, t+1} = V_{C \rightarrow U, t} + \Delta V_{C \rightarrow U, t}$$

Here we allow the representation to be formed by learning from the environment over the lifetime. To do this, we allow learning to take place over four epochs, each with 500 learning events. Each learning event randomly samples a relationship from the environment represented as an edge in the environment network. Then one of the associates is randomly assigned as the cue and the other as the outcome, and the representation is updated according to the Rescorla-Wagner model with $\alpha = 1$ and $\beta = .2$ and $\lambda_i = 1$.

Behaviour

The two stylized facts associated with cognitive aging are rising entropy and a reduction in pairwise similarity judgments. Each of these is recovered from the representation as follows.

Rising entropy. Rising entropy refers to the reduction in the predictability of free association targets as individuals age (Dubossarsky et al., 2017). We can measure this using *Shannon’s information entropy*. This measures the surprisingness of associate given the presence of a cue. Because the output of Rescorla-Wagner learning is a weighted edge, we can compute this for every cue in the network representation as follows:

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

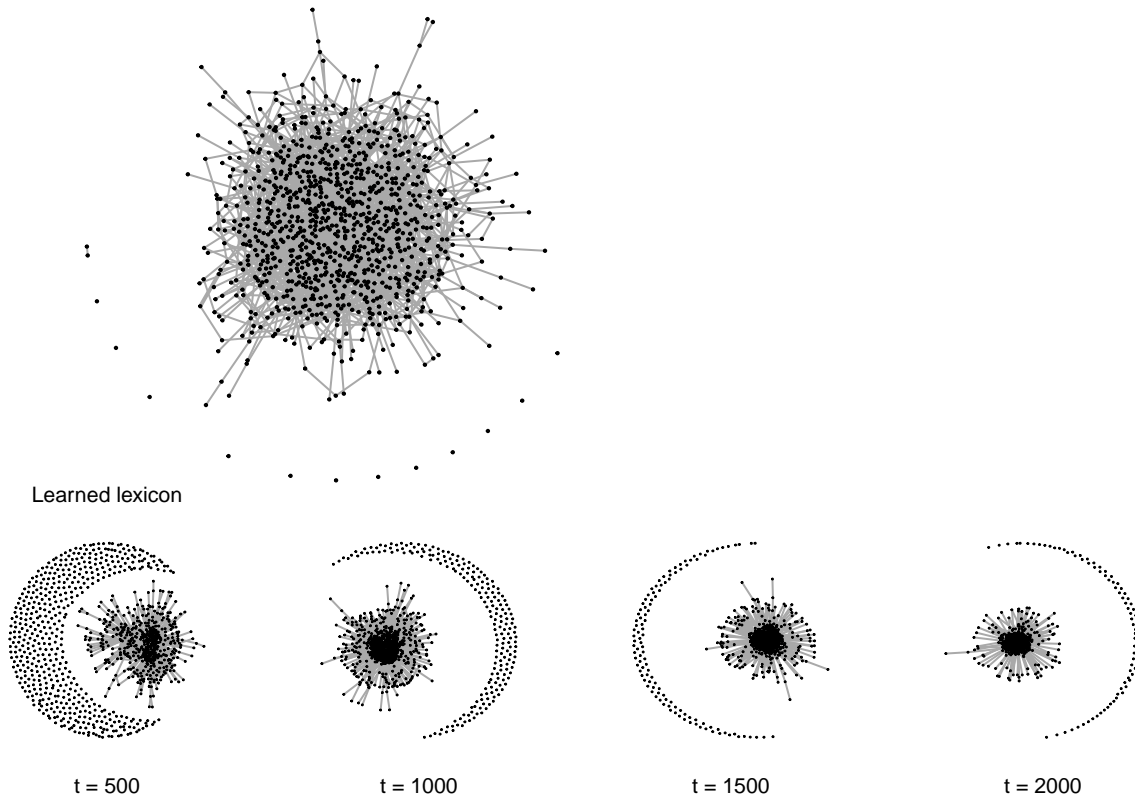
Here, p_i is the proportion of the weight along edge i for all k edges. That is, $p_i = \frac{w_i}{\sum_k w_k}$.

Similarity. To simulate similarity judgments, we will create a measure of co-activation between cues. To do this, we will allowing spreading activation to leave one node and measuring activation at the other node, $A_{j \rightarrow k}$. 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.

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

We will measure this similarity for all node pairs in the representation.

Results



Discussion

Because edges were formed in Dubossarsky et al. (2017) and Wulff et al. (2019) by requiring a minimum number of cue-target associations in the free association task, rising entropy corresponds to a less predictable pattern of cue-target associations and the lower likelihood of an edge between any cue-target pair. As a consequence, representational networks with higher entropy will produce behaviour networks that are more sparse.

References

- Amer, T., Wynn, J. S., & Hasher, L. (2022). Cluttered memory representations shape cognition in old age. *Trends in Cognitive Sciences*, 26(3), 255–267.
- 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.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Elsevier.
- 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.
- 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.
- 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.

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.

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

Petersen, A. M., Tenenbaum, J. N., Havlin, S., Stanley, H. E., & Perc, M. (2012).

Languages cool as they expand: Allometric scaling and the decreasing need for new words. *Scientific Reports*, 2(943), 1–10.

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.

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.

Serrano, M. Á., Flammini, A., & Menczer, F. (2009). Modeling statistical properties of written text. *PLoS ONE*, 4(4), e5372.

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

- 231 Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- 232 Trimmer, P. C., McNamara, J. M., Houston, A. I., & Marshall, J. A. (2012). Does natural
233 selection favour the rescorla–wagner rule? *Journal of Theoretical Biology*, 302, 39–52.
- 234 Wulff, D. U., De Deyne, S., Aeschbach, S., & Mata, R. (2022). Using network science to
235 understand the aging lexicon: Linking individuals’ experience, semantic networks, and
236 cognitive performance. *Topics in Cognitive Science*, 14(1), 93–110.
- 237 Wulff, D. U., De Deyne, S., Jones, M. N., & Mata, R. (2019). New perspectives on the
238 aging lexicon. *Trends in Cognitive Sciences*, 23(8), 686–698.
- 239 Zortea, M., Menegola, B., Villavicencio, A., & Salles, J. F. de. (2014). Graph analysis of
240 semantic word association among children, adults, and the elderly. *Psicologia: Reflexao
241 e Critica*, 27, 90–99.

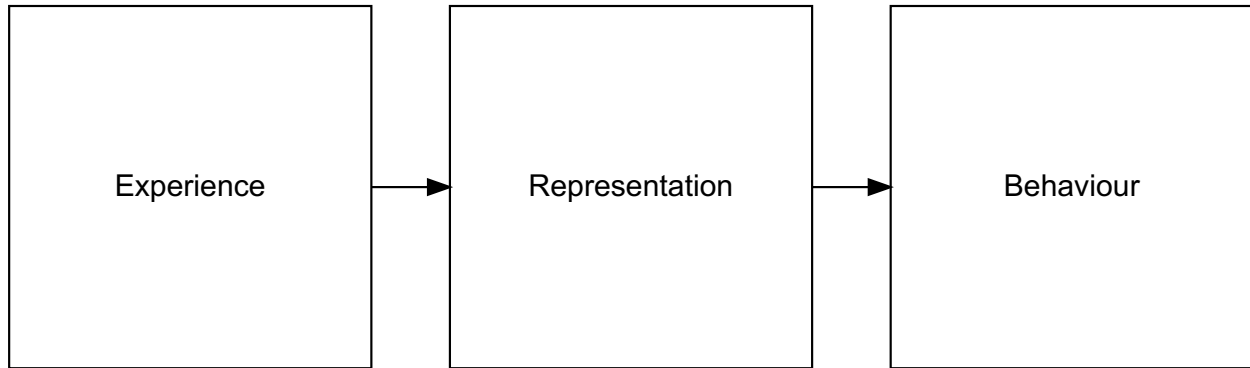


Figure 1. The process of translating experiences into behaviour via representation. Arrows represent processes that translate one domain into another. Thus learning translates experience into a representation and additional cognitive processes, such as associative recall, or similarity evaluations translate representations into behaviour.

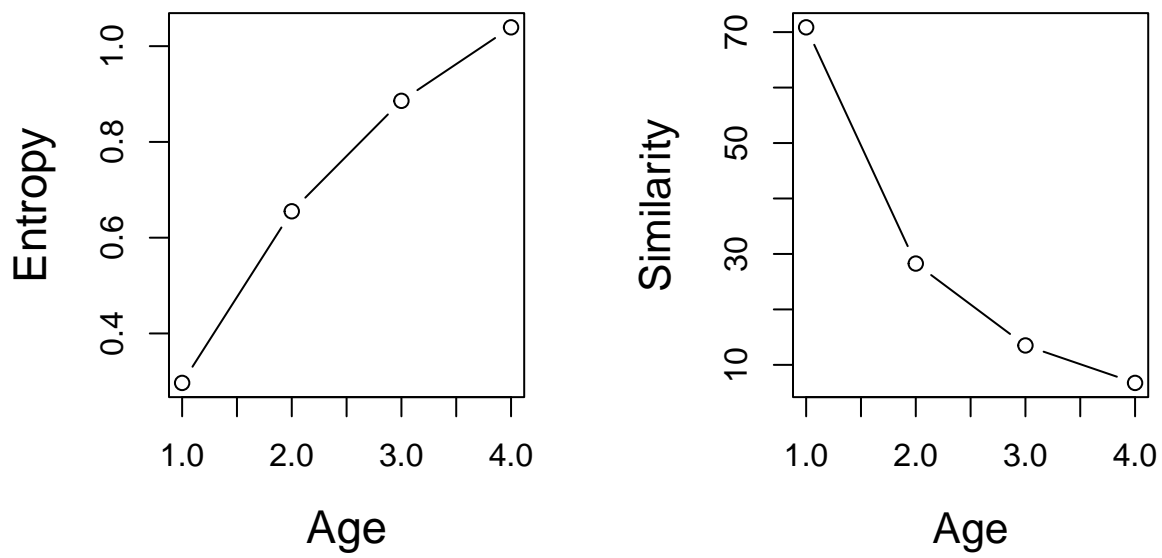


Figure 2. The rising entropy and falling similarity of the aging lexicon as a function of learning.