An enrichment account of cognitive aging

Thomas 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.

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

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

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-document 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), and the differences between them stereotypically distinguish the old from the young.

Several recent accounts of cognitive aging have argued for a relationship between these two phenomenon (Amer, Wynn, & Hasher, 2022; Buchler & Reder, 2007; e.g., Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014). That is, that a decline in fluid intelligence is the natural consequence of a rise in crystallized intelligence. This is proposed to arise either because new experiences violate prior learned expectations (e.g., Ramscar et al., 2014) or because prior experiences 'clutter' knowledge in a way that limits the speed with which old knowledge can be accessed.

Offering a formal account of this interactionist theory is the aim of the present article. However, it is useful to first juxtapose this theory against an often proposed alternative that treats fluid and crystallized intelligence 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) illustrates how this biological process 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). Moreover, neuropathology—associated with posthumously verified evidence of Alzheimer's, non-Alzheimer's neurodegenerative disease, and cerebrovascular conditions—can account for up to 40% of the variation in late-life cognitive decline (Boyle et al., 2021). This leaves substantial variance in cognitive-decline unexplained.

In addition, percent volume of grey-matter declines from early life and white-matter volume rises and then falls after the fifth decade (Ge et al., 2002; Giorgio et al., 2010). Cortical thickness also declines alongside concomitant increases in cerebrospinal fluid space (Lemaitre et al., 2012). These findings are usually associated with the word "atrophy" and fit with our intuition for biological aging.

There are many cognitive and brain related changes that are consistent with both accounts. For example, brain activity changes across the lifespan in relation to encoding and task processing, showing increased contributions from the default-mode network (Grady, Springer, Hongwanishkul, McIntosh, & Winocur, 2006). This phenomenology is also associated with decreased modularity within brain regions combined with larger interconnectivity between regions in later life (Geerligs, Renken, Saliasi, Maurits, & Lorist, 2015; Spreng & Turner, 2019). A reliance on exploiting past experience over exploration of novel environments may be an adaptive explanation for these changes (Spreng & Turner, 2021).

However, without a formal account of how these changes facilitate cognitive aging at the computational level, it is challenging to know what we should expect of cognitive aging, what is decline associated with atrophy, and what is simply decline we might expect of any system that learns.

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. Their proposal is that 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. Ramscar et al. (2014) took a different approach to explain age-related declines in paired-associate learning. They based their work on Desrosiers and Ivison (1986) observation that older adults perform more poorly on paired-associate learning than younger adults mainly on unrelated pairs—that is, on paired associations that would be least likely to be learned from prior experience. Ramscar, 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 Ramscar 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, producng 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, without understanding what we might expect from normal cognitive aging, efforts to explain cognitive aging as the result of degradation may attempt to bridge an explanatory gap that does not exist. Moreover, they may even get the causation backwards. Which is to say, if learning gives rise to some of the primary markers of

age-related cognitive decline, then so-called atrophy—while apparent—is unnecessary. As demonstrated below, by 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 learning relationships from the environment to develop a cognitive representation, and then using this representation to generate behavior. This involves three components:

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 are 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 Erdös-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:

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

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

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

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, $A_{j \to k}$. 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.

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

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 minimun 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 behavioural networks that are more sparse.

Write some more stuff.

Do people who learn more showing earlier degradation? look at clutter paper.

Ramscar 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,

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

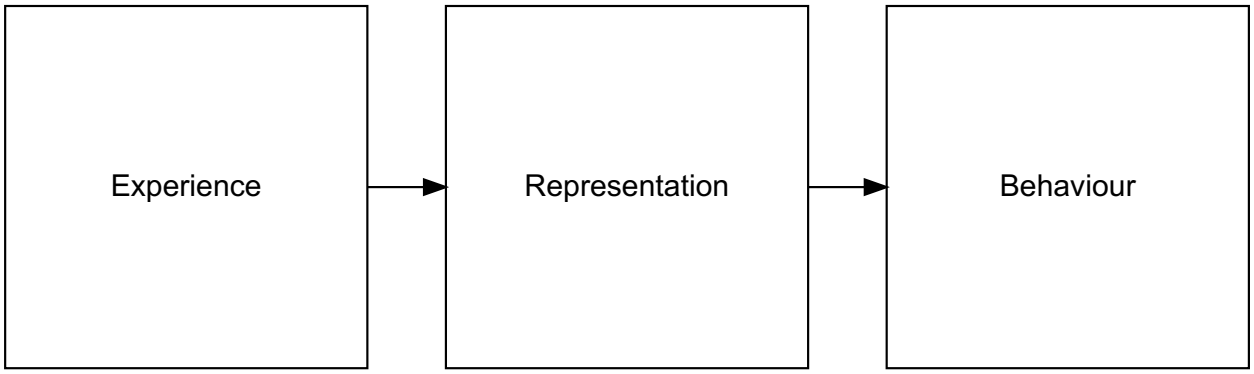

Figure 1. The process of translating experiences into behaviour via reprepresentation. Arrows represent processes that translate one domain into another. Thus learning translates experience into a representation and additional cognitive processes, such as associative recall, or similarity evaluations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