Learning and Cognitive Maturation

Michael Ramscar & Harald Baayen
Eberhard Karls Universität Tübingen
michael.ramscar@uni-tuebingen.de, harald.baayen@uni-tuebingen.de

Word Count
[2658, Includes Main Text, SEE ALSO: (cross-references), References and Further Readings.

Abstract
Healthy ageing is associated with cognitive changes that are commonly thought to reflect diminished processing power in our minds and brains. Formally, however, it is impossible to establish whether cognitive processing capacities actually do decline over the lifespan in the absence of functional models of cognitive processing, and without controlling for the way that learning tends to result in increases in the amount of information processed by the cognitive system. When the cognitive processes measured by psychometric tests of behavior are formally defined, the performance of older and younger adults can be seen to show little evidence of decline; rather, changes in test scores closely reflect the predictable performance of a system with consistent capacity processing increasing information loads as age and experience increase.

Main Text

It is widely believed that adult aging is inevitably accompanied by a progressive loss of cognitive function. This belief is apparently confirmed by the fact that adults' ages increase, their reaction times on many tasks get slower, and their memories, especially for names, appear to fade (Salthouse, 2009). Recently however, a number of findings have challenged conventional wisdoms. The idea that aging is necessarily coupled with declines in cognitive capacity has come under scientific challenge, as indeed has the notion that the neural structures that implement mental functions simply "atrophy" in our later years (Ramscar et al, 2013, 2014; Burke & Barnes, 2006).

Seen from the perspective of our attempts to understand and model ordinary, healthy cognitive functioning across the lifespan, the idea that age-related changes in reaction time and memory performance should necessarily be seen as signs of declines in performance or capacity runs counter to another, widely-accepted conventional wisdom, namely that our minds and brains can be best understood by treating them as a natural information processing device. It turns out that a simple fact about information processing systems is that speed and accuracy are, by definition, a function of the *amount* of information that is being processed and the capacity of the processing device (Shannon, 1948). Put simply, if processing capacity is held constant, and the amount of information that has to be processed is increased, something has to give.

Information processing systems are digital, not just in the commonly understood sense that they make use of binary codes of ones and zeros, but also in the more interesting sense that in information theory, the "information" that is communicated in systems is broken down into a set of discrete, discriminable states that can be encoded by various combinations of ones and zeros.

The amount of information in a system is then a function of the number of discrete states that can be discriminated within it, along with the way that these states are organized (Shannon, 1948). What makes this particularly relevant to human information processing – and to our understanding of cognitive aging – is that by far the best, and most detailed theories of how our minds / brains learn that we have also are also couched in discriminative terms. Although it may seem counterintuitive, a range of findings in psychology and neuroscience support the view that way in which we "add" new items to our memories is best characterized in terms of a process that increases the number states that our minds are capable of discriminating (Ramscar et al, 2010).

A natural consequence of this is that where learning serves to increase the number of discriminable experiences in an adult's memory, the amount of information that must be processed in order to access those memories will ultimately tend to increase, even allowing for the fact that practice may lead to gains in tests of performance for specific memory items (Ramscar et al, 2014). In the still-developing brains of children and adolescents, it is likely that some of the increase in the information load that is accrued from learning is offset by changes in neural morphology that effectively increase processing capacity (Fernald, Marchman, & Weisleder, 2012). However, given that the morphological structure of adults brains is relatively stable (Burke & Barnes, 2006), it follows that unless the information gains that are likely to result from experience are controlled for, it will be impossible to determine whether the changes in performance observed on psychometric tests as adults age are evidence of declining cognitive capacities or predictable consequences of learning on information processing.

To illustrate the importance of these points when it comes to understanding cognitive maturation, consider the age-related increase in the rates at which adults respond when they are asked to discriminate real words from non-words in tests. Although it is well known that older adults are slower to respond than younger adults, until recently researchers have tended to ignore the relationship between age and accuracy in the lexical recognition task. An analysis of The English Lexicon Project (Balota et al, 1999), a large dataset collected in order to allow analysis of lexical processing to be conducted on standard data, revealed that although older adults and younger adults are equally accurate when it comes to recognizing words that are frequently encountered in English, as the frequency of to-be-recognized words decreases, the accuracy of the young subjects accuracy plummets to approximately chance performance. By contest, while the old participants also make more errors for low frequency than high frequency words, they still outperform the young participants by a wide margin, getting 80% of their responses correct where the young subjects are at chance.

This analysis makes two things clear: First, older adults' increased ability to recognize rare words is a function of their much larger vocabularies (Brysbaert et al, 2014), which are in turn a product if their greater linguistic experience (in fact, such is the continuous nature of language learning across the lifespan that the diversity of the words speakers use reliably increases with age in even the briefest of exchanges; Meylan & Gahl, 2014); And second, some, if not all of, the change in older adults' response speeds will reflect the fact that checking to see whether a word is (or is not) part of a large vocabulary necessarily involves more searching than checking for a word in a small vocabulary. Indeed when bilinguals take these tests, the increased costs of having to search two vocabularies also increases response times, such that the speed of younger bilinguals is comparable to older monolinguals (Gollan, Montoya, Cera, & Sandoval, 2008).

Comparisons of bilinguals and monolinguals make it clear that slower reaction times in lexical decision tasks are not, in themselves, evidence of cognitive decline. Rather, what is clear is that becoming a bilingual *changes* the nature of the lexical decision task in a way that makes it harder. This same is true when older monolinguals are compared to their younger counterparts: When a younger adult decides that a low-frequency word she doesn't know is a "non-word," she makes what is to her an accurate and correct response, and this means that in a standard lexical decision task, young adults are be confronted with materials that contain – for them – a higher proportion of non-words than words. This will make real words more noticeable to younger adults, providing yet another reason for why it is a mistake to assume that response speeds straightforwardly reflect processing speed, or that slower response speeds are evidence of decline. Rather, as these considerations make clear, it makes little sense to compare two peoples' "processing speeds" on a task unless one has some idea of *what* they are processing, *how* that processing occurs and *how much* processing is involved for each individual. A comparison of "processing speeds" in the absence of a model of processing is analogous to a comparison of strength that considers only size of the objects people are carrying tells but not the weight of the objects, and will inevitably promote the same kind of false conclusion.

While a proper analysis of lexical decision data across the lifespan makes it clear that language learning is a lifelong process, the discriminative processes that characterize learning are best illustrated by considering another measure of cognitive performance, paired associate learning (PAL), a canonical measure of an individual's capacity to learn and recall new information. In PAL tests, people learn word pairs. Some are easy, i.e., *baby-cries*; others harder, *obey-eagle*. People perform worse on this task as they get older, a finding widely held to reveal considerable declines in cognitive function across the lifespan (Salthouse, 2009).

Quantitative analyses of PAL tests show that traditional interpretations of PAL results paint a misleading picture of our cognitive abilities, first because they fail to take into account the fact that knowledge about the words being tested changes with age and experience, and second because it has been known for more than half a century that prior knowledge exerts considerable influence on the associative learning of stimulus-response (S-R) pairs probed by PAL tests. To explain why it is worth reviewing what we have learned about learning since the Russian physiologist Ivan Pavlov conducted the famous experiments in which he conditioned dogs to salivate at the sound of a bell. Pavlov's initial results led to a view of learning called associationism: if a cue is present, and an outcome follows, it is widely believed that animals will inevitably learn to associate them. While humans also often learn this way, the word "associate" is misleading in this context. This process is actually a result of a more complex process that implicitly tests and refines the predictions that prior experience supports. These implicit predictions serve to determine which features of the environment are reliable indicators of the events that can be expected to unfold.

Three main factors have been shown to determine the degree to which cues and responses become associated in learning (Rescorla, 1988): the frequency at which a cue co-occurs with an outcome (the *association rate*) tends to promote learning, such that high association rates tend lead to strong associations, whereas the frequency with which the cue occurs without an outcome (the *cue background rate*) and the degree to which an outcome is predictable in the presence of a cue (learning is *blocked* learning when outcome predictability is high). Tracking the the information provided by these factors ensures that successful predictors are reinforced, at the expense of unreliable predictors, which our brains learnt to ignore, such that a dog associates a bell with food only because it has learned to ignore all the other cues available to it (Ramscar et al, 2010).

The discriminative logic that characterizes our formal understanding of the learning processes involved in "associative learning" leads in turn to a revised view of the way that changes in PAL performance should be interpreted: Across all ages, adults find some pairs, such as *baby-cries* easier than other pairs, such as *obey-eagle*, and over the lifespan, while PAL learning gets harder for all pairs as we grow older, harder word-pairs become proportionally more difficult to learn. Why? The explanation for why some pairs are easy is straightforward. We know, through experience, that babies cry. When given the word pair *baby-cries*, it is easy to remember to produce the word 'cries' when prompted with the word 'baby.' By contrast, experiences involving *obeying* hardly ever coincide with, or are followed by, experiences with *eagles*. Accordingly, as our experience of situations that involve *obeying* but not *eagles* accumulates over the lifetime, we will become increasingly confident situations involving obeying will not lead to eagles. This means the an older adult instructed to learn a pair like *obey-eagle* in a test is being asked to remember a pairing that experience increasingly teaches us is nonsensical, and that will be increasingly actively inhibited in the ordinary course of learning from that experience.

Or, to put it another way, nonsensical pairings between words make increasingly less sense the more experience one accumulates over the lifetime. Indeed from a learning perspective, if learning processes continue to function across the lifespan, we *ought* to expect the PAL task become harder with age. Until recently, however, it would have been impossible to determine just how much of the changes seen in PAL performance across the lifespan might be attributed to learning. There was simply no way of measuring how differences in experience in something as complex as language might influence learning in something as apparently simple as a PAL task. Today, however, computational analyses enable us to estimate the connections between words based on their patterns of occurring together in billions of words of English text and speech. Using these co-occurrence frequencies as a proxy for the co-occurrence frequencies of events and experiences, a multiple regression analysis of PAL performance in Ramscar et al (2013) showed that adults' sensitivity to the patterns of word co-occurrence that can be expected to determine the learning of PAL pairs increases over the lifetime. Older speakers have stronger positive beta weights for the actual co-occurrence frequency of words pairs than younger speakers, and stronger negative beta weights for background rates, and the patterns of co-occurrence that cause blocking. Indeed, the analysis revealed as healthy adults get older, the degree to which they find individual PAL pairs easy or hard to learn increasingly reflects the actual differences in difficulty that the predicted by the word co-occurrence patterns of English.

When analyzed in terms of learning, and with a detailed characterization of the linguistic environment in which learning takes place, adults' PAL performance reveals: First, that language learning continues across the lifespan; and Second, that the increased sensitivity to the sensibility of pairings shown by older adults is far from evidence of cognitive decline, since it actually reflects a better understanding of the information structure of the English lexicon (Ramscar et al, 2013).

Similar patterns of results have been obtained for other traditional cognitive measures in which performance has been taken to support the idea that cognitive processing declines with age (i.e., letter classification, and memory for names), whilst perhaps more strikingly, the same analyses that explain why older adults scores *decline* on PAL tests can also explain why they *improve* on measures of verbal fluency: in both cases, changing scores reflect the changes in the structure of the information being processed that result from experience, and not changes in information processing or cognitive function (Ramscar et al, 2014).

The findings reviewed here offer an alternative perspective that challenges many long held assumptions. They suggest that when the cognitive processes measured by psychometric tests of behavior are formally defined, the performance of older and younger adults show little evidence of decline, and indicate that changes in test scores reflect changes in performance in cognitive systems that are predicted by the increased information loads that are inevitably accrued from a lifetime of learning.

SEE ALSO:
[Include cross-references here. See the cross-references list of other entries on the homepage of ScholarOne.]

References and Further Readings

Balota, D. A., Cortese, M. J., & Pilotti, M. (1999). Item-level analyses of lexical decision performance: Results from a mega-study. Abstracts of the 40th Annual Meeting of the Psychonomics Society (P. 44). Los Angels, CA: Psychonomic Society.

Brysbaert, M., Keuleers, E., Mandera, P. & Stevens, M (2014) The first results of the Groot Nationaal Onderzoek Woordenschat. Presentation at the 24th Computational Linguistics in the Netherlands Conference, Leiden, Netherlands, January 17th, 2014

Burke, S. & Barnes, C. (2006). Neural plasticity in the aging brain. Nature Reviews Neuroscience, 7 (1): 30–40

Fernald, A., Marchman, V. A. & Weisleder, A. (2012). SES differences in language processing skill and vocabulary are evident at 18 months. Developmental Science,16(2), 234-248.

Gollan, T. H., Montoya, R. I., Cera, C. M., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language, 58, 787–814.

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

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

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The Effects of Feature-Label-Order and their implications for symbolic learning. Cognitive Science, 34(6), 909-957.

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

Salthouse, T. A. (2009). When does age-related cognitive decline begin?. Neurobiology of aging, 30(4), 507-514.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379–423.

Author Mini-Biography:

[**Michael Ramscar** holds degrees in Philosophy, Information Technology Systems and Artificial Intelligence. He is a senior research scientist at Tuebingen Univeristy, where **Harald Baayen** is the William Von Humboldlt Professor of Quantitative Linguistics.]

Key Words

[Include list from ScholarOne.]