

Aging, Neuropsychology, and Cognition

A Journal on Normal and Dysfunctional Development

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/nanc20>

Relevance of working memory for reinforcement learning in older adults varies with timescale of learning

Irene van de Vijver & Romain Ligneul

To cite this article: Irene van de Vijver & Romain Ligneul (2020) Relevance of working memory for reinforcement learning in older adults varies with timescale of learning, *Aging, Neuropsychology, and Cognition*, 27:5, 654-676, DOI: [10.1080/13825585.2019.1664389](https://doi.org/10.1080/13825585.2019.1664389)

To link to this article: <https://doi.org/10.1080/13825585.2019.1664389>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 22 Sep 2019.



[Submit your article to this journal](#)



Article views: 1196



[View related articles](#)



[View Crossmark data](#)



OPEN ACCESS



Relevance of working memory for reinforcement learning in older adults varies with timescale of learning

Irene van de Vijver^{a,b,c} and Romain Ligneul^d

^aBehavioural Science Institute, Radboud University, Nijmegen, The Netherlands; ^bDepartment of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands; ^cAmsterdam Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands; ^dChampalimaud Neuroscience Program, Champalimaud Foundation, Lisboa, Portugal

ABSTRACT

In young adults, individual differences in working memory (WM) contribute to reinforcement learning (RL). Age-related RL changes, however, are mostly attributed to decreased reward prediction-error (RPE) signaling. Here, we investigated the contribution of WM to RL in young (18–35) and older (≥ 65) adults. Because WM supports maintenance across a limited timescale, we only expected a relation between RL and WM with short delays between stimulus repetitions. Our results demonstrated better learning with short than long delays. A week later, however, long-delay associations were remembered better. Computational modeling corroborated that during learning, WM was more engaged by young adults in the short-delay condition than in any other age-condition combination. Crucially, both model-derived and neuropsychological assessments of WM predicted short-delay learning in older adults, who further benefitted from using self-conceived learning strategies. Thus, depending on the timescale of learning, age-related RL changes may not only reflect decreased RPE signaling but also WM decline.

ARTICLE HISTORY

Received 9 October 2018

Accepted 2 September 2019

KEYWORDS


Aging; reinforcement learning; working memory; individual differences; computational modeling

Introduction

No matter how old we are, we need to keep updating our behavior to successfully interact with our environment. Such behavioral updating requires learning from the consequences of our actions, or reinforcement learning (RL). However, RL can occur on different timescales: the interval between consecutive learning experiences may vary from milliseconds to years. For example, the second bite of a newly tried fruit will follow the first one very rapidly if the first bite was evaluated positively. Selecting the destination of your next vacation, however, may only happen once or twice a year. Thus, constant rainfall at the last location will only influence a new decision months later. The impact of this temporal variability on the cognitive and neural processes supporting RL has long been overlooked in the field and completely ignored in aging studies.

A crucial mechanism underlying RL is the integration of outcomes over learning experiences into “cached values”, as described by the principles of Rescorla and

CONTACT Irene van de Vijver  irenevandevijver@gmail.com

 Supplemental data for this article can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Wagner (1972) or TD-learning rules (Sutton & Barto, 1998). More specifically, a dopaminergic signal reflecting a reward prediction error (RPE) is thought to progressively and automatically strengthen or weaken the associations between context- or stimuli-related cues and performed actions (Frank, 2005; Frank, Seeberger, & O'reilly, 2004; Glimcher, 2011; Schultz, 2013; Schultz, Dayan, & Montague, 1997). Highly dependent upon the activity of the basal ganglia, this type of learning is slow and incremental, but relatively robust to brief fluctuations in outcome contingencies.

However, the implementation of RL by the human brain is not limited to RPE-based learning within the basal ganglia. It competes and interacts with other processes such as model-based learning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) and working memory (WM; Baddeley & Hitch, 1974) which both engage prefrontal and parietal cortices. In particular, working memory has recently been shown to support stimulus-response association learning (Collins, Brown, Gold, Waltz, & Frank, 2014; Collins & Frank, 2012). As associations can be stored in WM instantaneously, WM allows for faster learning and more flexible updating of these associations than lower-level, RPE-based learning. Yet, WM is temporally limited (Baddeley, 2012), so its involvement likely varies with the temporal scale of learning. Indeed, the influence of WM on RL depends on both the delay between stimulus repetitions and the WM load imposed by task demands (Collins & Frank, 2012). Additionally, the relative contribution of RPE-based learning versus WM learning seems to vary with WM capacity: individuals with higher model-derived WM capacity estimates can continue to use WM up to larger delays between stimulus repetitions than individuals with lower WM capacity, who tend to rely more on RPE-based learning already with smaller delays (Collins, Ciullo, Frank, & Badre, 2017).

Behavioral updating based on previous experiences declines with age: older adults show increased error rates and may never reach the same level of accuracy as young adults (see, e.g., Eppinger & Kray, 2011; Hämmerer, Li, Müller, & Lindenberger, 2011; Mell et al., 2005; Simon, Howard, & Howard, 2010; van de Vijver, Ridderinkhof, et al., 2016; Weiler, Bellebaum, & Daum, 2008). The extent of this decline depends both on specific characteristics of the aging individual and on task demands (Eppinger, Hämmerer, & Li, 2011; Hämmerer & Eppinger, 2012; Ligneul, 2019; van de Vijver, Ridderinkhof, et al., 2015).

Age-related changes in RL have been attributed to changes in striatal and dopaminergic RPE signaling (Chowdhury et al., 2013; Eppinger et al., 2011; Eppinger, Schuck, Nystrom, & Cohen, 2013). Indeed, with age, gray-matter density in the striatum decreases (Raz et al., 2003; Tisserand et al., 2004), and dopamine receptor and transporter availabilities reduce (Karrer, Josef, Mata, Morris, & Samanez-Larkin, 2017). However, age-related changes extend beyond these systems. One of the brain regions the most severely affected by aging is the frontal cortex (see, e.g., Bennett, Madden, Vaidya, Howard, & Howard, 2010; Burzynska et al., 2010; Raz et al., 2005; Salat et al., 2009). This decline affects a range of cognitive functions, including WM and executive functioning, which are also dependent on DA signaling (Burzynska et al., 2012; Charlton, Barrick, Lawes, Markus, & Morris, 2010; Cools & D'Esposito, 2011; Grieve, Williams, Paul, Clark, & Gordon, 2007; Madden et al., 2010; Ziegler et al., 2010).

If WM is involved in and benefits RL in older adults in a similar way as in young adults, age-related WM decline should also affect RL in this age group, so that age-related changes in RL may depend not only on decreased RPE signaling, but also on changes in WM. Supporting this hypothesis, individual WM capacity in older (as well as young) adults predicts stimulus-response-outcome association learning (van de Vijver, Ridderinkhof, et al., 2015). Given the temporal limitations of WM, this relation should mostly hold when

stimulus-response associations need to be stored only briefly until the same stimulus reappears. Thus, the first goal of the current study was to investigate whether RL relates to individual WM capacity in older adults when delays between learning experiences are short.

The use of WM to support RL may not only influence how associations between behavior and outcomes are learned, but also how they are stored for later use. RPE-based learning is slow and incremental, but stored associations are very robust and can even result in inflexible, automatically triggered habits (Dickinson, 1985; Packard & Knowlton, 2002). However, the use of WM for RL has been demonstrated to coincide with a decrease in striatal RPE signals, thereby likely decreasing consolidation of learned associations (Collins et al., 2017; Wittmann et al., 2005). Therefore, the second goal of this study was to investigate whether consolidation of learned associations was better for learning experiences separated by a longer temporal delay, in which WM would not play a significant role. Since the involvement of WM in learning from experiences separated by a short delay is hypothesized to depend on individual WM capacity, this could also affect RPE signals and consolidation. We therefore additionally explored whether consolidation of associations learned with short delays was related to individual differences in WM.

To answer these questions, 34 young and 35 older adults were invited to perform a deterministic RL task in which they had to learn stimulus-response associations by trial-and-error using the feedback they received. Crucially, the presentations of half of the stimuli were concentrated within short periods of time (short-delay condition), whereas the presentations of the other half of the stimuli were distributed over the whole task (long-delay condition). Knowledge of the learned associations was tested immediately after learning and a week later. To verify the relative involvement of RPE-based learning and WM in the two conditions and age groups, we additionally fitted the computational model combining RPE-based learning and WM as developed by Collins and Frank (Collins & Frank, 2012) to the choices of participants during the learning phase.

Our first hypothesis was that learning performance would be lower in older compared to young adults. Second, since the additional recruitment of WM in the short-delay condition was expected to aid learning (Collins et al., 2017; Collins & Frank, 2012; van de Vijver, Ridderinkhof, et al., 2015), we further hypothesized that learning would be faster in the short-delay than in the long-delay condition, and that this effect would be more salient in young adults due to an age-related decline in WM capacity. Third, because the use of WM was expected to lead to decreased RPEs, performance on the second test was hypothesized to be better in the long-delay compared to the short-delay condition. Finally, we expected learning in the short-delay condition, but not in the long-delay condition, to correlate with individual WM scores.

Methods

Participants

Thirty-four young and 35 older adults participated in this study. The data of 5 young and 5 older adults were excluded, because they used psychotropic medication (1 young), were non-native Dutch speakers (4 young), or performed an incorrect version of the learning task (5 older). The remaining 29 young adults (6 male, 5 left-handed) ranged in age from 19 to 29 years

(M 22.41, SD 2.86), the 30 older adults (14 male, 3 left-handed) from 65 to 79 years (M 70.13, SD 3.95). All participants were free of neurological and psychiatric disorders, did not take psychotropic medication, did not have a history of brain damage, and had a normal or corrected-to-normal vision. Older adults did not show signs of depression as assessed with the Geriatric Depression Scale 15 (score of 5 or lower; Burke, Roccaforte, & Wengel, 1991; Yesavage et al., 1982), or of severe cognitive decline as assessed with the Cognitive Screening Test 20 (score of 17 or higher; Deelman, Maring, & Otten, 1989; Ponds, Verhey, Rozendaal, Jolles, & Deelman, 1992). Participants received course credits or financial compensation for their participation. The amount of compensation was fixed and not related to performance on any of the tasks. This study was approved by the Ethics Committee Social Science of the Radboud University.

Reinforcement learning task

Participants were required to learn associations between multiple pictorial stimuli and two response buttons (keyboard keys “z” and “/”) by trial-and-error, using the feedback they received after each choice. On each trial, a picture was presented until the participant pressed a button (max 2000 ms), followed by feedback (1200 ms; Figure 1(a)). Trials were separated by a 2000 ms inter-trial interval (ITI). Stimuli were colored versions of pictures from the Snodgrass set (Snodgrass & Vanderwart, 1980), representing, e.g., animals, food, and tools. Feedback consisted of a smiling (positive) or a sad (negative) face icon. If participants did not respond within 2000 ms, feedback consisted of the words “TE LAAT” (too late).

The task included 32 stimuli, which were each presented 16 times. These 512 trials were divided into four blocks of 128 trials. Crucially, half of the stimuli featured in the short-delay condition, the other half in the long-delay condition (Figure 1(b)). The stimuli in the long-delay condition were presented four times in every block. In addition, per block four stimuli from the short-delay condition were presented 16 times each. Thus, each block consisted of 4 presentations of all 16 long-delay stimuli and 16 presentations of 4 short-delay stimuli. Four new short-delay condition stimuli were selected for each block. Trial order was randomized per 32 trials, including one presentation of each long-delay stimulus and four presentations of the four short-delay condition stimuli. The resulting average distance between stimulus repetitions was 7.70 trials (SD 5.89, range 1–45) in the short-delay condition, and 32.03 trials (SD 13.05, range 1–63) in the long-delay condition.

Participants received extensive instructions and started with a practice block containing task-unrelated stimuli. In this block, 2 short-delay stimuli were presented 15 times each, and 2 long-delay stimuli 5 times each. After each block, participants were informed of the number of points they collected: every correct/incorrect button press resulted in the gain/loss of one point.

Test of learned associations

The test phase following the learning task was similar but did not include feedback (Figure 1(b)). This phase consisted of 64 trials, including two presentations of each stimulus. Presentation order was randomized per complete stimulus set (32 trials). After participants pressed one of the two buttons (no time limit) and an ITI of 2000 ms, the next picture was presented. When they finished the test, participants were informed of how many correct answers they had provided.

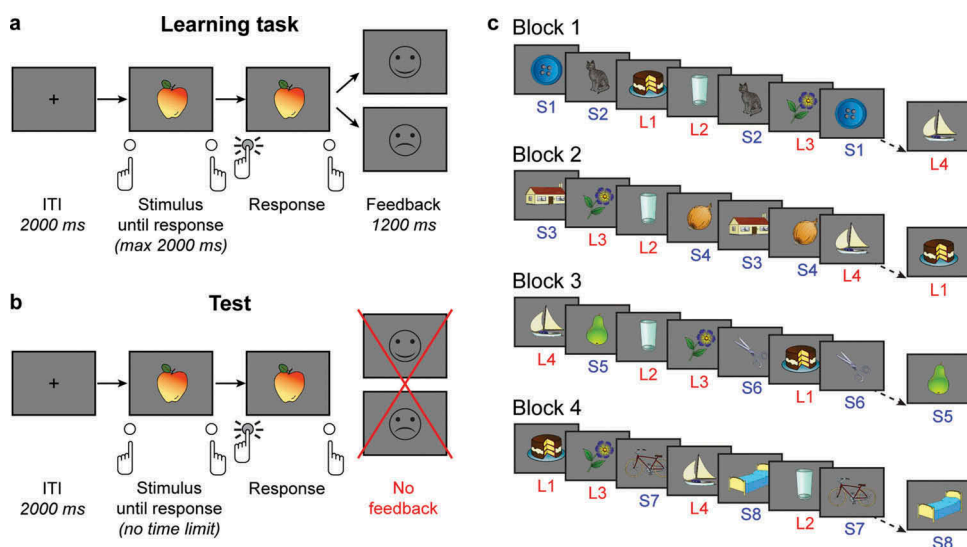


Figure 1. Reinforcement learning task and test of learned stimulus-response associations. (a) Sequence of events in an example trial of the RL task. (b) Sequence of events in an example trial of the test phase. The test phase was identical to the learning phase except for: 1. the absence of a response deadline, and 2. the absence of performance feedback. (c) Illustration of the trial sequences in the RL task. Each stimulus represents a complete trial. Each block featured stimuli from both the short- and the long-delay (S and L) condition. Note that the figure is a simplification. In the real task, there were 32 stimuli in each condition. Stimuli in the long-delay condition were presented 4 times in each block, stimuli in the short-delay condition were presented 16 times in one block.

Neuropsychological measures

WM was assessed with the Operation Span test (O-span, Turner & Engle, 1989). In this test, participants have to remember words while performing mathematical operations. The number of words and operations per set varied between 2 and 6. Performance was scored using the partial credit scoring system (Conway et al., 2005): percentages correctly supplied words at the correct location in a set were first averaged over sets with the same length, and subsequently over all set lengths. Fluid intelligence was assessed with the Matrix Reasoning test from the Wechsler Adult Intelligence Scales IV (WAIS IV), an abstract problem-solving test (Wechsler, 2008). Crystallized intelligence was assessed with the Nederlandse Leestest voor Volwassenen (NLV, Dutch reading test; Schmand, Bakker, Saan, & Louman, 1991).

Procedure

Participants performed two sessions that took place one week apart. The first session consisted of the RL task and the test of the learned associations. In this session, a final question asked whether participants used a learning strategy (yes/no), and if so, to describe the strategy. The second session started with a second performance of the test of the learned associations, followed by the O-span, the Sensitivity to Punishment and Sensitivity to Reward questionnaire (Torrubia, Ávila, Moltó, & Caseras, 2001; not reported

here), the WAIS Matrix Reasoning test and finally the NLV. Both sessions ended with a brief questionnaire about the participant's motivation and experiences.

Statistical analyses of behavior

Accuracy on the RL task was defined as the percentage correct button presses on the trials that participants responded to. Accuracy scores were entered into a mixed ANOVA with within-subject factors Delay condition (short, long) and Bin (stimulus presentation 1–4, 5–8, 9–12, 13–16), and between-subject factor Age (young, older). Note that in the short-delay condition, stimuli were presented 16 times within one block, so all four stimulus presentation bins were part of the same block. In the long-delay condition, however, because the stimuli were presented only four times within each block, each of the four stimulus presentation bins was in fact part of a separate task block. Because we did not have specific hypotheses about the effects of the two delay conditions on reaction times (RTs), the same analysis with RTs as dependent variable is reported in the Supplementary material.

Accuracy on the tests was defined as the percentage correct button presses. Test accuracy scores were entered into a mixed ANOVA with within-subject factor Test (1, 2), and between-subject factor Age (young, older). Because test accuracy scores immediately after learning were almost perfect in both age groups, the effects of condition were only assessed for the second test. Test 2 accuracy scores were entered into a mixed ANOVA with within-subject factor Delay condition (short, long) and between-subject factor Age (young, older). We did not examine RTs during the test because we did not impose a response deadline in this part.

In the short-delay condition, four new stimuli were used in each block. However, knowledge of the associations from the long-delay condition increased over blocks, which may have affected the processing of the short-delay stimuli. Because both learning and recall of the associations in the short-delay condition may therefore have differed depending on the block they appeared in, additional analyses examining the effect of block on accuracy and RTs are presented in the Supplementary material.

To investigate whether performance during the learning task and subsequent tests depended on individual WM capacity, per age group we computed Spearman correlations between O-span scores and learning performance in both delay conditions (average percentage correct over all bins and blocks), and between O-span scores and test performance on the second test for short-delay condition stimuli (percentage correct). For all ANOVAs, Greenhouse-Geisser corrections were applied when required, but uncorrected degrees of freedom are reported.

Computational modeling

We used a computational model developed by Collins and Frank (2012) to account for the choices of participants in the task. As its name indicates, the “Reinforcement-Learning and Working Memory” (RLWM) model arbitrates between two distinct learning modules competing to explain choices.

The reinforcement-learning module uses a classical Rescorla-Wagner rule to update values in response to the feedback delivered in each trial:

$$Q_{t+1}^{RL}(S) = Q_t^{RL}(S) + \alpha * (O_t - Q_t^{RL}(S))$$

Here, $Q_{t+1}^{RL}(S)$ corresponds to the predicted value of a left choice for the stimulus S ($S \in \{1, 2, \dots, 32\}$). Thus, the outcome variable O_t is arbitrarily coded as 1 whenever the left button was pressed and a positive feedback was received, or when the right button was pressed and a negative feedback was received. Otherwise, O_t was coded as 0. Therefore, $Q_t^{RL}(S)$ converged toward 1 at a pace defined by the learning rate α for the 16 stimuli associated with a left response and toward 0 for the 16 stimuli associated with a right response. Q was initialized ($t = 1$) at a value of 0.5, reflecting the lack of prior knowledge about the correct responses.

The reinforcement learning module further included a perseveration mechanism such that $\alpha = \alpha(1 - pers)$ whenever a negative feedback was given ($pers \in [0, 1]$). Thus, a *pers* parameter of 1 implies a complete absence of update in response to negative feedback, whereas a *pers* parameter of 0 implies a symmetrical update in correct and incorrect trials.

By contrast, the working memory module fully updated the value of responding left after each feedback O_t (coded in the same way as above), so that:

$$Q_{t+1}^{WM}(S) = O_t$$

Importantly, both modules included a decay mechanism implemented at the beginning of each trial, so that values progressively returned to the indifference starting point of 0.5 at a rate controlled by the decay parameters Φ_{RL} and Φ_{WM} ($\Phi \in [0, 1]$):

$$Q_t = Q_t + \phi * (0.5 - Q_t)$$

Note that, unlike Collins and Frank model, we did not explicitly model working memory capacity. Indeed, because we did not systematically vary the total number of stimuli to be learned from one block to another, our task did not allow the estimation of this parameter. This was the only difference between the model used here and the model by Collins and Frank. Instead, working memory capacity was implicitly reflected in the decay parameter of the WM module and in the arbitration parameter responsible for the weighting of Q_t^{WM} and Q_t^{RL} at the decision stage:

$$Q_t^{decision}(S) = \rho_{WM} * Q_t^{WM}(S) + (1 - \rho_{WM}) * Q_t^{RL}(S)$$

The arbitration parameter ρ_{WM} ($\rho_{WM} \in [0, 1]$) played a crucial role for our study because it reflected the variable contribution of the working memory module from one condition to another (i.e., short *versus* long) and from one group to another (i.e., young *versus* old). Finally, the composite value $Q_t^{decision}(S)$ was turned into a probability of choosing the left button using a sigmoid transform:

$$P(left, S) = \frac{\varepsilon}{2} + (1 - \varepsilon) * \frac{1}{1 + e^{-\beta * Q_t^{decision}(S)}}$$

The β parameter corresponds to the inverse temperature determining to which extent participants' choices were driven by the composite values $Q_t^{decision}$ ($\beta \in \mathbb{R}$), with higher values of β reflecting more consistent decisions (note that all beta were found to be positive). Finally, the random noise parameter ε ($\varepsilon \in [0, 1]$) determined to which

extent participants engaged in value-independent exploration (i.e., exploration unconstrained by learned values).

Model fitting was performed using the Variational Bayesian Analysis (VBA) toolbox (Daunizeau, Friston, & Kiebel, 2009). Compared to non-Bayesian methods, this tool has the advantage of accounting for the uncertainty related to estimated model parameters and of informing the optimization algorithm about prior distributions of parameters' values. To the exception of β (mean 0 and variance 50), all priors were initially defined as Gaussian distribution of mean 0 and variance 3. Indeed, such initialization approximates the uniform distribution over the [0–1] interval following sigmoid transformation. Thus, all parameters except β were sigmoid-transformed to restrain their variation to the [0–1] interval. After fitting the data using these uninformative priors, we used the distributions obtained from this first-pass as priors for a second-pass. Indeed, using empirical priors helps to stabilize the fits and implies a more stringent constraint on between-group statistical tests since the same set of prior distributions is applied to all participants.

Results

Age-related differences in general cognitive functions

Age-related differences in WM and intelligence followed normal developmental patterns (see Table 1): young adults outperformed older adults on both the O-span, $t(57) = 5.846$, $p < .001$, $d = 1.917$, and the WAIS Matrix Reasoning test, $t(51.2) = 5.323$, $p < .001$, $d = 1.383$. When WAIS Matrix Reasoning performance was corrected for age using norm scores, group differences disappeared, $t(57) = 0.697$, $p = .489$, $d = .183$, suggesting that differences between groups could indeed be attributed mainly to age. As expected, older compared to young adults scored higher on the NLV, $t(57) = 3.543$, $p = .001$, $d = .926$.

Reinforcement learning with short versus long delays

Participants responded on almost all trials, with a maximum of 1 missed trial per condition in young adults ($M = 0.138$, $SD = 0.351$ missed trials in both delay conditions) and 6 missed trials per condition in older adults (short: $M = 1.400$, $SD = 1.604$ missed trials, long: $M = 1.000$, $SD = 1.156$ missed trials). Although the number of missed trials is significantly higher in older compared to young adults (short: $t(34.301) = -3.847$, $p < .001$; long: $t(31.7670) = -4.139$, $p < .001$), it is important to note that both groups missed less than 1% of trials. Learning accuracy was lower in older compared to young adults, $F(1,57) = 34.385$, $p < .001$, $\eta_p^2 = .376$ (Figure 2(a)). Participants were able to learn the correct associations: accuracy increased over trial bins, $F(3,55) = 602.816$, $p < .001$, $\eta_p^2 = .914$, although the duration of this increase differed

Table 1. Scores of young and older adults on measures of intelligence and WM (WAIS MR = WAIS Matrix Reasoning test).

Measure	Young				Old			
	<i>M</i>	<i>SD</i>	Min	Max	<i>M</i>	<i>SD</i>	Min	Max
WAIS MR raw	21.24	3.18	13	25	15.73	4.65	7	24
WAIS MR norm	11.83	2.78	6	16	11.33	2.67	6	17
NLV	82.31	4.68	74	91	87.77	6.90	72	98
O-Span	71.73	9.69	50.22	96.67	48.52	14.12	9.00	72.00

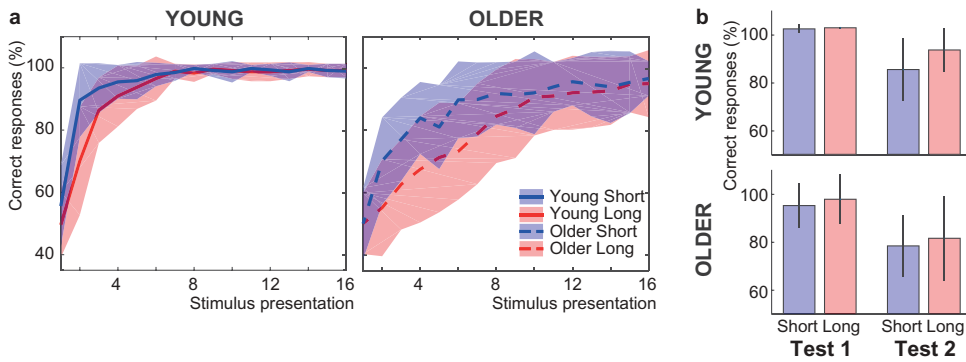


Figure 2. Behavioral performance of young and older adults in the reinforcement learning task. (a) Response accuracy per stimulus presentation in the short- and long-delay condition as the percentage of correct responses over stimuli. Response accuracy was lower in older compared to young adults. Both groups showed higher response accuracy in the short- compared to the long-delay condition (colored patches represent standard deviations). (b) Recall of the learned associations in the short- and long-delay condition immediately after learning (Test 1) and a week later (Test 2) as the percentage of correct responses over stimuli. Recall of the learned associations was worse in older than in young adults, and worse during Test 2 compared to Test 1. During Test 2, recall was better for associations learned in the long-delay condition in both age groups (error bars represent SD).

between age groups, $F(3,55) = 28.639$, $p < .001$, $\eta_p^2 = .334$; accuracy increased between all bins in older adults (all p -values $< .001$), whereas it stabilized after the third bin in young adults (bin 1 – bin 2, bin 2 – bin 3: p -values $< .001$, bin 3 – bin 4: $t(28) = -0.278$, $p = .783$).

Accuracy was higher in the short-delay than in the long-delay condition, $F(1,57) = 53.016$, $p < .001$, $\eta_p^2 = .482$. The difference between conditions decreased with learning, $F(3,55) = 28.444$, $p < .001$, $\eta_p^2 = .333$, and disappeared in the fourth bin (bins 1 and 2: p -values $< .001$, bin 3: $t(58) = -2.217$, $p = .031$, bin 4: $t(58) = -1.068$, $p = .290$). The learning benefit in the short-delay condition was larger in older compared to young adults, $F(1,57) = 10.397$, $p = .002$, $\eta_p^2 = .154$, but it was significant in both age groups (young: $t(28) = 4.234$, $p < .001$, $d = .786$; older: $t(29) = 6.058$, $p < .001$, $d = 1.106$). This effect was further specified by a three-way interaction, $F(3,55) = 5.584$, $p = .005$, $\eta_p^2 = .089$, which indicated that young adults only learned better in short-delay than in the long-delay condition in the first bin (bin 1: $t(28) = -5.338$, $p < .001$, bins 2–4: p -values $> .15$), likely because they reached ceiling afterward, whereas the difference between conditions remained present until the third bin in the older adults (bin 1–2: p -values $< .001$, bin 3: $t(29) = -2.194$, $p = .036$, bin 4: $t(29) = -1.190$, $p = .244$).

To summarize, older adults had more difficulty learning the correct associations than young adults did. Both age groups learned better in the short-delay condition.

Consolidation of associations after learning with short versus long delays

Older adults performed worse than young adults did on the tests of the learned associations, $F(1,57) = 14.971$, $p < .001$, $\eta_p^2 = .208$ (Figure 3(a)). Both groups remembered the correct associations better immediately after learning than a week later, $F(1,57) = 126.703$, $p < .001$, $\eta_p^2 = .690$, but the performance was above chance level in both age groups during both tests (all p -values < 0.001). The decrease in overall recall accuracy between test sessions did not differ between age groups, $F(1,57) = 1.865$, $p = .177$, $\eta_p^2 = .032$. In line

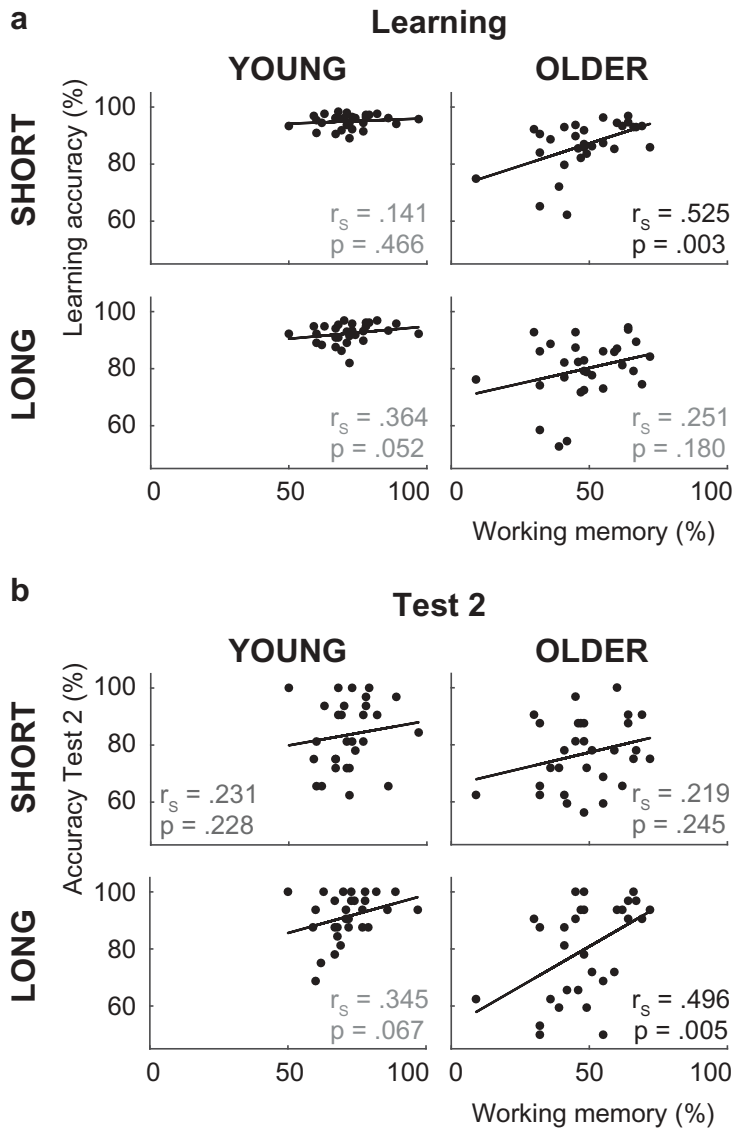


Figure 3. Spearman correlations between individual differences in working memory capacity and behavioral performance in (a) the learning phase and (b) Test 2 of the reinforcement learning task. In older adults, individual differences in WM capacity correlated with learning in the short-delay condition. Individual differences in WM were not related to consolidation. Note that removing the older adult with the lowest WM score does not affect the pattern of significant correlations in either the learning or the test phases.

with the general age effect, older adults also performed worse on Test 2, $F(1,57) = 10.159$, $p = .002$, $\eta_p^2 = .151$. A main effect of condition indicated that during Test 2, recall was better for long- than short-delay associations, $F(1,57) = 10.659$, $p = .002$, $\eta_p^2 = .158$. This effect of delay condition did not differ between age groups, $F(1,57) = 2.059$, $p = 0.157$, $\eta_p^2 = 0.035$.

To summarize, recall of the learned associations was worse in older than in young adults. A week after learning, recall was better for associations learned in the long-delay condition in both age groups.

Relation between task performance and individual differences in working memory

In older adults, O-span scores correlated with learning performance in the short-delay condition, $r_s = .525$, $p = .003$ (Figure 3(a)), but not the long-delay condition, $r_s = .251$, $p = .180$. A direct comparison of these correlations (Steiger, 1980) indicated that they differed significantly, $z = 2.004$, $p = .045$, two-tailed. In young adults, O-span scores did not correlate with learning performance in either condition (short: $r_s = .141$, $p = .466$, long: $r_s = .364$, $p = .052$). Because the absence of significant correlations in the young adults could be due to a ceiling effect on the learning task in this group and a resulting lack of variability in total accuracy scores, we subsequently examined Spearman correlations in both delay conditions for the four trial bins separately (trials 1–4, 5–8, 9–12, 13–16). In the young adults, there were still no significant correlations in either delay condition in any bin. In the older adults, the correlation between O-span scores and learning was significant in the first two bins (see Table 2 for all correlations). Interestingly, in the last bin O-span scores of older adults correlated with performance in both learning conditions (note, however, that no correction was applied for multiple tests).

A week after learning, test performance for associations learned in the short-delay condition did not correlate with O-span scores in either age group (older adults: $r_s = .219$, $p = .245$, young adults: $r_s = .231$, $p = .228$; Figure 3(b)).

To sum up, learning behavior was related to individual differences in WM in older adults. In this age-group, WM correlated with learning in the short-delay condition. Individual differences in WM were not related to consolidation of associations learned in the short-delay condition.

Table 2. Correlations between O-span scores and learning in the short- and long-delay conditions per trial bin (bins 1–4 include trials 1–4, 5–8, 9–12, and 13–16, respectively).

			r	p
Young adults	Short delay	Bin 1	0.214	0.265
		Bin 2	−0.035	0.855
		Bin 3	−0.047	0.809
		Bin 4	0.065	0.739
	Long delay	Bin 1	0.362	0.054
		Bin 2	0.342	0.069
		Bin 3	−0.231	0.227
		Bin 4	−0.166	0.391
Older adults	Short delay	Bin 1	0.507	0.004**
		Bin 2	0.528	0.003**
		Bin 3	0.261	0.164
		Bin 4	0.388	0.034*
	Long delay	Bin 1	0.196	0.299
		Bin 2	0.149	0.431
		Bin 3	0.275	0.141
		Bin 4	0.399	0.029*

*Significant at $\alpha = 0.05$; **significant at $\alpha = 0.01$.

Computational model estimation of the role of working memory in reinforcement learning

The computational model served two primary functions in the context of our study: (i) validating our hypothesis that performance depended more on working memory in the short-delay condition than in the long-delay condition; (ii) confirming that older adults relied on average more on an RPE-based reinforcement-learning process than young adults did.

Thus, we fitted the computational model to the choice data of the short- and long-delay conditions *separately* and we examined the two sets of estimated values for the RLWM arbitration parameter in the two age groups. While the arbitration parameters did not differ between age groups, $F(1,57) = 1.756$, $p = .190$, $\eta_p^2 = .030$, we observed a salient age-condition interaction, $F(1,57) = 10.199$, $p = .002$, $\eta_p^2 = .152$ (Figure 4(a)). The arbitration parameter was similar across conditions in older adults, $t(29) = .170$, $p = .866$ (long = 0.28 ± 0.12 ; short = 0.27 ± 0.11), but it differed in young adults, who relied more on their working memory in the short as compared to the long condition, $t(28) = -4.827$, $p < .001$ (long = 0.24 ± 0.11 ; short = 0.37 ± 0.13). In the short condition specifically, the arbitration parameter of young and older adults differed significantly ($t(57) = 3.04$, $p = 0.003$). Still, performance in this condition was predicted by WM decay in both age groups (young: $r_s = -.596$, $p = .001$; old: $-.805$, $p < .001$).

In a more exploratory vein, we examined the relationship between O-span and model parameters, and we investigated age differences in model parameters other than arbitration. To do so, we used the parameter values obtained when fitting the model to the complete dataset. Interestingly, O-span scores correlated with perseveration in the young but not in the older group (young: $r_s = -.43$, $p = .019$; old: $r_s = -.10$, $p = .590$; $z = 1$), whereas it correlated with WM decay in the older but not in the young group (young: $r_s = -.170$, $p = .36$; old: $r_s = -.362$, $p = .049$). Note however that these correlation coefficients did not differ significantly across groups, all $p > 0.1$, hence requiring confirmation in a larger sample.

Finally, it must be noted that within the parameters governing the update and maintenance of stimulus-response associations only the decay rates of the RL module were unaffected by age per se, $p = 0.24$ (Figure 4(c)): the learning rate of the RL module was higher in young as compared to older adults, $t(57) = 5.23$, $p < .001$, whereas perseveration following negative feedback (RL module) was higher in older adults, $t(57) = 4.83$, $p < 0.001$, and so was WM decay, $t(57) = 5.64$, $p < .001$. By contrast, the parameters governing the decision process (i.e., random noise and inverse temperature) did not differ between age groups, all $p > 0.2$.

Together these results indicate that aging alters the updating and maintenance of values but leaves decision-making essentially intact. Moreover, the pattern of correlations observed with the O-span scores suggests that the detrimental effect of lower WM functioning for performance might be to some extent mediated by distinct pathways in older (i.e., WM decay) *versus* young (i.e., perseveration) adults.

Benefits of using a learning strategy

Twenty out of the 30 older and 24 out of the 29 young adults reported using a learning strategy. The actual strategies varied, but many involved grouping the stimuli that were related to the same response button into meaningful collections (for the participant),

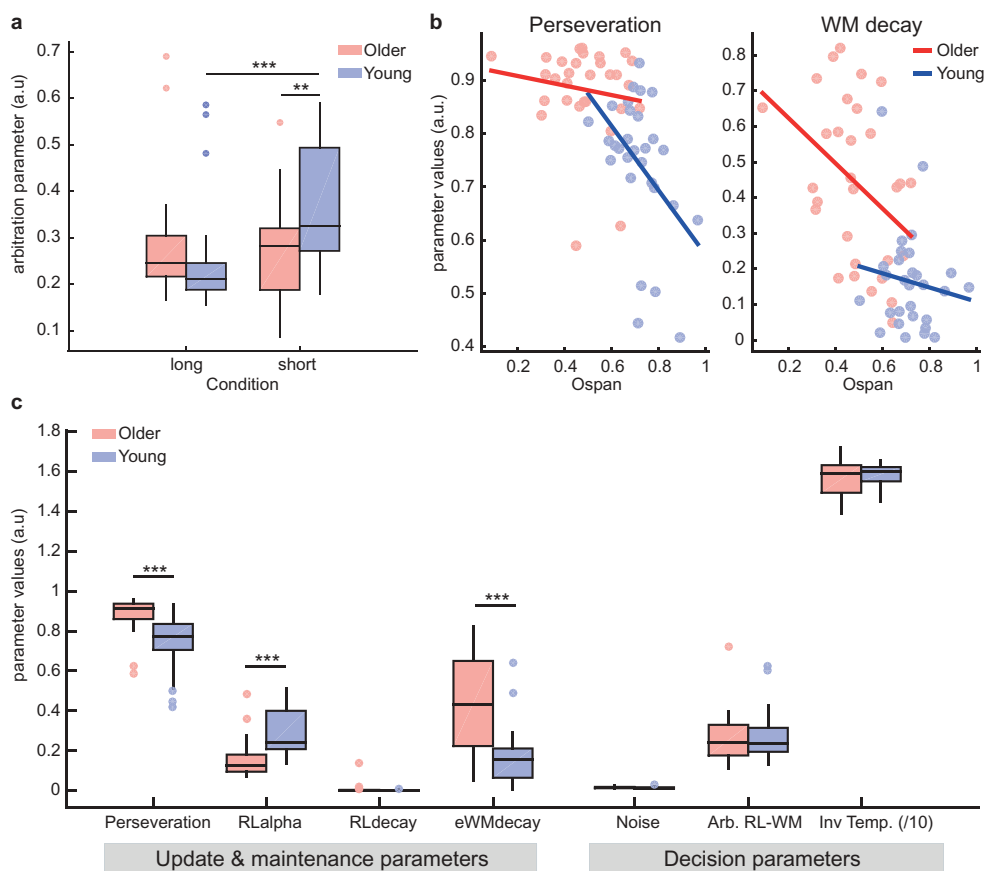


Figure 4. Computational modeling results. (a) The analysis of the arbitration parameter showed that young adults relied more on the working memory module in the short-delay condition, as compared to older adults and to the long-delay condition. (b) Two model parameters correlated strongly with O-span scores in the separate age groups: the decay rate of the associations stored in the WM module in older adults, and perseveration in front of negative feedbacks in young adults. (c) These parameters, as well as the learning rate of the RL (i.e., RPE-based) module, were also significantly different in older as compared to young participants. Overall, older participants showed significantly higher WM decay and perseveration, as well as significantly lower learning rates within the RL module (Arb. = arbitrator, Inv Temp. = inverse temperature; * $p < .05$, ** $p < .01$, *** $p < .001$).

and explicitly storing the exceptions (for example: “all fruits go left, except for the bananas”). Grouping could be based on, for example, the similarity in the content of the pictures, specific visual features, or relations between pictures. Because such chunking can help keep larger amounts of information available in WM (Cowan, 2010; Miller, 1956), we explored the effect of using a strategy on learning performance, and the relation with WM capacity. Because the number of young adults not reporting using a strategy was very small ($n = 5$), we limited these comparisons to the older adults.

Learning and test accuracy for older adults that did and did not report using a strategy are presented in Table 3. A comparison of these two groups indicated that strategy use benefitted performance both during learning and subsequent recall of

Table 3. Learning accuracy and recall of the learned associations in older adults that did and did not use a learning strategy.

Task part	Strategy (<i>n</i> = 20)		No strategy (<i>n</i> = 10)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Learning Short	89.56	4.76	81.13	11.81
Learning Long	84.02	7.25	71.74	11.83
Test 1 Short	95.31	7.28	88.13	8.69
Test 1 Long	98.91	2.33	88.44	14.21
Test 2 Short	80.78	11.44	69.38	9.52
Test 2 Long	83.59	15.46	72.81	17.05

the learned associations. Strategy-using learners had higher learning accuracy in the short-delay, $t(10.486) = 2.171$, $p = .054$, $d = .937$ (marginally significant) as well as the long-delay condition, $t(12.493) = 3.012$, $p = .010$, $d = 1.252$. Recall was higher on test 1 for the short-delay, $t(28) = 2.392$, $p = .024$, $d = .897$, and the long-delay stimuli, $t(9.243) = 2.314$, $p = .045$, $d = 1.028$. Recall was also higher on test 2 for the short-delay condition, $t(28) = 2.712$, $p = .011$, $d = 1.084$, but not for the long-delay condition, $t(28) = 1.740$, $p = .093$, $d = .662$. Whereas WM scores were numerically higher in participants that reported using a strategy ($M = 51.55\%$, $SD = 12.55$) than in participants that did not ($M = 42.30\%$, $SD = 15.66$), this difference was not statistically significant, $t(28) = 1.752$, $p = .091$, $d = .652$.

Discussion

The aim of this study on age-related changes in RL was two-fold: to investigate how the role of WM during learning differs with the temporal delay between RL experiences, and to examine the impact of this delay on subsequent recall of the learned associations. In line with our hypotheses, we found better learning when repetitions of learning experiences were separated by short compared to long delays. Crucially, in older adults only learning with short delays was predicted by individual WM capacity. Fitting a computational model combining WM- and RPE-based learning supported the relevance of WM in coping with task demands, especially in the short-delay condition. Additionally, it indicated specific deficits in the update and the maintenance of stimulus-response associations but not in decision-making in older adults, and confirmed that older adults who showed a lesser WM decay performed closer to young adults. Consolidation of learned information was better when learning experiences were separated by longer delays. We also explored the spontaneous use of self-conceived learning strategies, the application of which seems to have benefitted both learning and recall in older adults.

The role of working memory in instrumental learning

Although studying the RPE-based learning system has proven useful for understanding the basic mechanisms underlying RL, the limitations of this relatively simple model in capturing RL in daily life have also been acknowledged (Doll, Simon, & Daw, 2012; Gershman & Daw, 2017; Wimmer & Shohamy, 2012). This realization has inspired research into interactions between RPE-based learning and other memory systems,

including WM (Collins et al., 2017; Collins & Frank, 2012) and episodic memory (Gershman & Daw, 2017; Wimmer, Daw, & Shohamy, 2012; Wimmer & Shohamy, 2012). In young adults, the involvement of WM in simple stimulus-response learning has previously been demonstrated (Collins et al., 2014; Collins & Frank, 2012). Collins and colleagues proposed a hybrid computational model including both RPE-based learning mechanisms and WM to account for their collaboration during RL (Collins & Frank, 2012). Combining this model with fMRI, they were able to demonstrate that the use of WM for learning coincided with a reduction in striatal RPE signaling. Individual differences in (model-derived) WM capacity predicted when participants switched from WM-based to RPE-based learning with increasing task loads and delays between stimulus repetitions, such that participants with higher capacities were able to rely on WM longer with increasing loads and delays (Collins et al., 2017).

The current behavioral results corroborate the idea that the relation between WM and RL depends on the delay between learning experiences, and demonstrate that this relation continues into older age: there was a learning advantage in the short- compared to the long-delay condition in both age groups, and individual differences in WM scores predicted learning in older adults when learning experiences followed each other shortly. The current correlation between learning and individual WM capacity in older adults resembles the correlation that we previously demonstrated in a stimulus-response-outcome learning task similar to the short-delay condition in the current study (van de Vijver, Ridderinkhof, et al., 2015). Relatedly, older adults have recently been demonstrated to employ a smaller focus of attention during RL, and to be less able to simultaneously evaluate multiple sources of information and extract the relevant information for learning, another aspect that may rely on WM (Radulescu, Daniel, & Niv, 2016).

In order to better understand the variable contribution of WM in the two task conditions in each age group, we fitted the computational model developed by Collins and Frank (2012) to the choice data of the training phase. This approach corroborated our hypothesis that the short-delay condition tapped more into WM than the long-delay condition did. Indeed, when both conditions were fitted separately, we observed that the arbitration parameter controlling the relative weights given to RPE-based and WM-based modules more strongly favored WM use for short- as compared to long-delay condition stimuli, especially in young adults. The successful recruitment of WM processes might thus explain the ceiling performances observed in that condition for most young adults.

Interestingly, age had no direct impact on decision-release parameters: the inverse temperature controlling value-driven choice and random noise controlling value-independent exploration were both similar in young and older adults. By contrast, age affected most of the parameters responsible for the updating or maintenance of stimulus-response associations: The increased rates of WM decay and the increased perseverative tendencies observed in older adults are in line with previous reports of age-dependent increases in WM decay rates (Ligneul, 2019) and perseverative tendencies (Head, Kennedy, Rodrigue, & Raz, 2009; Ridderinkhof, Span, & van der Molen, 2002), and concur to suggest that age differences in performance might result from a lesser engagement of the prefrontal executive system.

Finally, it is interesting to note that the learning rates of the RPE-based module were lower in older as compared to young adults. While our hypothesis and most of our

results point toward WM-dependent deficits, it is thus possible that healthy aging also exerts a more discrete but significant effect on lower-level processes such as DA signaling thought to control RPE-based learning rates (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Rutledge et al., 2009).

Even though the modeling results suggest that young adults were able to recruit WM in the short-delay condition, we did not replicate the relation between behavioral assessments of WM and RL that we previously demonstrated in young adults (Van De van de Vijver, Ridderinkhof, et al., 2015). Even in the first part of the learning curve (first bin), no relation between WM and RL was found in this age group. Since young adults reached an average accuracy level above 90% within only a few trials, the absence of a relation may have been caused by a ceiling effect and related lack of individual variability. Thus, although the relation between WM capacity and RL has previously been established in young adults, the question whether this relation also varies with the timescale of learning in this age group should be further investigated by adapting task difficulty to their abilities.

Our task design differed from previous studies in the delay between stimulus repetitions (Collins et al., 2017; Collins & Frank, 2012; van de Vijver, Ridderinkhof, et al., 2015): whereas previous studies used delays ranging from 1 to 6 trials, the average delay in our short-delay condition already exceeded this. Additionally, after the initial presentation of all long-delay stimuli during each block 20 associations had to be remembered and kept available, 4 short-delay plus 16 long-delay associations. Thus, the average delay in combination with the load in the current design likely exceeded the capacity limits of WM that have previously been put forward (Collins & Frank, 2012; Cowan, 2010). It is possible that participants in the current study have engaged a selection mechanism, focusing attention and thereby WM on a subset of the stimuli.

The current task design also deviated from previous research (Collins et al., 2017; Collins & Frank, 2012) by presenting the stimuli from the separate delay conditions intermixed. We chose this design in an attempt to equalize the influence of the continuous memory load on both delay conditions. However, whereas stimuli from both the short- and the long-delay condition were new in the first block, in the subsequent blocks, part of the long-delay associations were likely already stored. This may have facilitated the selection of the new short-delay condition associations for storage in WM.

Relatedly, it is important to note that in the current task increasing temporal delays between stimulus repetitions coincided with increasing numbers of intervening trials. This means that when delays between stimulus repetitions were longer, more of the other associations were updated and thus, more mental operations were performed. We therefore cannot distinguish whether the difference between conditions is purely depended on the temporal delay or rather on the cognitive performance required during that delay. Additionally, WM capacity and fluid intelligence are known to be related (Unsworth, Fukuda, Awh, & Vogel, 2014). It is therefore hard to disentangle the respective contribution of each in the facilitation of learning.

WM-related brain activity centers on lateral prefrontal cortex (PFC; Cohen et al., 1997; Curtis & D'Esposito, 2003; Rypma, Prabhakaran, Desmond, Glover, & Gabrieli, 1999), which is thought to support the control and organization of information maintenance (Bor, Duncan, Wiseman, & Owen, 2003). In line with the idea that the use of WM allows for faster learning than low-level RPE-based dopaminergic processes, IPFC is active

mainly during initial RL (Collins et al., 2017; Seger & Cincotta, 2006). However, the striatum and frontal cortex are part of a densely interconnected system (Di Martino et al., 2008; Haber & Knutson, 2010; Haruno & Kawato, 2006), and besides the basal ganglia, PFC is one of the main recipients of dopaminergic projections originating in the midbrain dopaminergic nuclei (Bannon & Roth, 1983). Indeed, age-related changes in dopaminergic processing are not limited to the striatum: a recent meta-analysis suggested similar age-related declines in dopaminergic receptor and transporter availability in the striatum, frontal cortex, and midbrain (Karrer et al., 2017). Recently, we have also demonstrated that different networks of frontostriatal connections are related to RL performance in older compared to young adults, including additional tracts between the striatum and lateral PFC (van de Vijver et al., 2016). Thus, further research including neuroimaging measures is needed to increase our understanding of age-related changes in the balance between brain networks related to WM and low-level, RPE-based learning to support RL.

Effects of stimulus spacing on consolidation

In both age groups, recall of the learned associations was almost perfect immediately after learning, and it was still high a week after learning. As predicted, consolidation was better for associations acquired in the long- than the short-delay condition. This is in line with the idea that WM use during acquisition in the short-delay condition coincides with decreased striatal RPE signaling (Collins et al., 2017), resulting in decreased consolidation (Wittmann et al., 2005). Note, however, that since learning was slower in the long-delay condition, participants likely experienced more RPEs in this condition than in the short-delay condition. Additionally, more surprising outcomes have been demonstrated to attract more attention and lead to larger RPEs, affecting both consolidation and learning about expected rewards (Ludvig, Madan, McMillan, Xu, & Spetch, 2018; Rouhani, Norman, & Niv, 2018). Given their relative infrequency, in the current experiment, the long-delay trials may have inspired more surprise and, thus, led to increased attention and RPE signaling. We therefore cannot rule out that the difference in recall between the delay conditions was not only caused by decreased RPE signals in the short-delay condition, but also by the larger number of RPEs or the increased attention and RPE signaling in the long-delay condition.

The memory benefit of more distributed spacing of stimulus presentations during learning has been extensively studied in the field of paired associate learning and it has also been observed in paradigms varying from episodic memory and motor learning to Pavlovian fear conditioning, although it is still unclear whether the underlying mechanisms are the same (Toppino & Gerbier, 2014). Most studies report comparable effects of distributed spacing on learning in young and older adults (Balota, Duchek, & Paullin, 1989; Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Logan & Balota, 2008; Maddox, Balota, Coane, & Duchek, 2011), although a more recent study testing recall 10 days after acquisition found a smaller effect of spaced presentations in older compared to young adults (Simone, Bell, & Cepeda, 2013). Many theories exist about the origin of this distributed-spacing effect, centering on three main ideas: memory is hindered by immediate repetition because processing of the second item is decreased if it is redundant with the first, memory is facilitated if a stimulus is encoded in multiple ways or contexts, and memory is facilitated if an item induces reactivation of its previous presentation, thereby strengthening the

previously encoded representation (Toppino & Gerbier, 2014). Thus, the distributed practice in the long-delay condition in our task may have affected long-term performance in multiple ways: it may have increased the impact of dopaminergic RPE-based fluctuations that strengthen consolidation, and also facilitated the storage in and recall from LTM with richer and more multifaceted representations.

Learning strategies

A majority of both young and older adults reported using a self-conceived learning strategy. Many strategies seem to have relied on the categorization of stimuli that were associated with the same response. Such structuring or “chunking” of information into associated combinations increases WM capacity as chunks rather than separate pieces of information can be stored (Cowan, 2010; Miller, 1956; Simon, 1974). Indeed, lateral PFC becomes more active when to-be-stored information can be combined into meaningful units (Bor et al., 2003). In the current experiment chunking likely took place in both conditions and therefore cannot explain the specific relation between WM and short-delay learning, but it may have contributed to increasing the amount of information that could be retained.

As we did not formalize our questions about learning strategies, we relied on the verbal descriptions of the participants. We therefore cannot reliably investigate the relation between specific types of strategies and learning or test performance, or whether older adults that show higher cognitive functioning are better able to come up with more appropriate strategies. Additionally, the number of older that reported not to have used a learning strategy was quite low for appropriate statistical analyses. Still, the role of self-conceived learning strategies in successful RL has so far received little attention in either young or older adults. Our results suggest that these strategies may play an important role in RL in older adults and may provide an interesting avenue for improving learning in this age group, although the relevance and specific role of such strategies need to be confirmed in more rigorous experiments.

Conclusions

The current study demonstrates that age-related changes in RL are not only due to the previously demonstrated decrease in dopaminergic RPE signaling (see, e.g., Chowdhury et al., 2013; Eppinger et al., 2011, 2013), but also related to WM declines. Additionally, it is a first indication of how the mechanisms supporting RL may vary depending on the timescale of learning. As older adults seem to use different frontostriatal brain networks than young adults do to process performance feedback and improve behavior (van de Vijver et al., 2016; van de Vijver, Cohen, & Ridderinkhof, 2014), the investigation of the interplay between cortical and subcortical changes and the related cognitive implications may provide promising insights into the dynamics of age-related changes in behavioral adaptation, and the differences therein with different timescales of learning.

Acknowledgments

The author would like to thank Sanne de Wit for helpful comments on a previous version of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29. doi:10.1146/annurev-psych-120710-100422.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York, NY: Academic. <http://books.google.com/books?hl=nl&id=o5LScJ9ecGUC&pgis=1>
- Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag, and retention interval. *Psychology and Aging*, 4(1), 3–9. Retrieved from <http://dx.doi.org/10.1037/0882-7974.4.1.3>
- Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging*, 21(1), 19–31. doi:10.1037/0882-7974.21.1.19
- Bannon, M. J., & Roth, R. H. (1983). Pharmacology of mesocortical dopamine neurons. *Pharmacological Reviews*, 35(1), 53–68. Retrieved from <http://pharmrev.aspetjournals.org/content/35/1/53.short>
- Bennett, I. J., Madden, D. J., Vaidya, C. J., Howard, D. V., & Howard, J. H., Jr. (2010). Age-related differences in multiple measures of white matter integrity: A diffusion tensor imaging study of healthy aging. *Human Brain Mapping*, 31(3), 378–390. doi:10.1002/hbm.20872
- Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, 37(2), 361–367. doi:10.1016/S0896-6273(02)01171-6
- Burke, W. J., Roccaforte, W. H., & Wengel, S. P. (1991). The short form of the geriatric depression scale: A comparison with the 30-item form. *Journal of Geriatric Psychiatry and Neurology*, 4(3), 173–178. doi:10.1177/089198879100400310
- Burzynska, A. Z., Nagel, I. E., Preuschhof, C., Gluth, S., Bäckman, L., Li, S.-C., ... Heekeren, H. R. (2012). Cortical thickness is linked to executive functioning in adulthood and aging. *Human Brain Mapping*, 33(7), 1607–1620. doi:10.1002/hbm.21311
- Burzynska, A. Z., Preuschhof, C., Bäckman, L., Nyberg, L., Li, S.-C., Lindenberger, U., & Heekeren, H. R. (2010). Age-related differences in white matter microstructure: Region-specific patterns of diffusivity. *NeuroImage*, 49(3), 2104–2112. doi:10.1016/j.neuroimage.2009.09.041
- Charlton, R. A., Barrick, T. R., Lawes, I. N. C., Markus, H. S., & Morris, R. G. (2010). White matter pathways associated with working memory in normal aging. *Cortex*, 46(4), 474–489. doi:10.1016/J.CORTEX.2009.07.005
- Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Düzel, E., & Dolan, R. J. (2013). Dopamine restores reward prediction errors in old age. *Nature Neuroscience*, 16(5), 648–653. doi:10.1038/nn.3364
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386(6625), 604–608. doi:10.1038/386604a0
- Collins, A. G. E., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *The Journal of Neuroscience*, 34(41), 13747–13756. doi:10.1523/JNEUROSCI.0989-14.2014
- Collins, A. G. E., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working memory load strengthens reward prediction errors. *The Journal of Neuroscience*, 37(16), 4332–4342. doi:10.1523/JNEUROSCI.2700-16.2017
- Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *The European Journal of Neuroscience*, 35(7), 1024–1035. doi:10.1111/j.1460-9568.2011.07980.x

- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. doi:[10.3758/BF03196772](https://doi.org/10.3758/BF03196772)
- Cools, R., & D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, 69(12), e113–e125. doi:[10.1016/j.biopsych.2011.03.028](https://doi.org/10.1016/j.biopsych.2011.03.028)
- Cowan, N. (2010). The magical mystery four. *Current Directions in Psychological Science*, 19(1), 51–57. doi:[10.1177/0963721409359277](https://doi.org/10.1177/0963721409359277)
- Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9), 415–423. doi:[10.1016/S1364-6613\(03\)00197-9](https://doi.org/10.1016/S1364-6613(03)00197-9)
- Daunizeau, J., Friston, K. J., & Kiebel, S. J. (2009). Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models. *Physica D: Nonlinear Phenomena*, 238(21), 2089–2118. doi:[10.1016/j.physd.2009.08.002](https://doi.org/10.1016/j.physd.2009.08.002)
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. doi:[10.1016/j.neuron.2011.02.027](https://doi.org/10.1016/j.neuron.2011.02.027)
- Deelman, B., Maring, W., & Otten, V. (1989). De CST, een gestandaardiseerde screeningsmethode voor dementie. In J. Schroots, A. Bouma, G. Braam, A. Groeneveld, D. Ringoir, & C. Tempelman (Eds.), *Gezond zijn is ouder worden* (pp. 163–170). Assen: Van Gorcum.
- Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., . . . Milham, M. P. (2008). Functional connectivity of human striatum: A resting state FMRI study. *Cerebral Cortex*, 18(12), 2735–2747. doi:[10.1093/cercor/bhn041](https://doi.org/10.1093/cercor/bhn041)
- Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 308(1135), 67–78. doi:[10.1098/rstb.1985.0010](https://doi.org/10.1098/rstb.1985.0010)
- Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, 22(6), 1075–1081. doi:[10.1016/J.CONB.2012.08.003](https://doi.org/10.1016/J.CONB.2012.08.003)
- Eppinger, B., Hämmerer, D., & Li, S.-C. (2011). Neuromodulation of reward-based learning and decision making in human aging. *Annals of the New York Academy of Sciences*, 1235, 1–17. doi:[10.1111/j.1749-6632.2011.06230.x](https://doi.org/10.1111/j.1749-6632.2011.06230.x)
- Eppinger, B., & Kray, J. (2011). To choose or to avoid: Age differences in learning from positive and negative feedback. *Journal of Cognitive Neuroscience*, 23(1), 41–52. doi:[10.1162/jocn.2009.21364](https://doi.org/10.1162/jocn.2009.21364)
- Eppinger, B., Schuck, N. W., Nystrom, L. E., & Cohen, J. D. (2013). Reduced striatal responses to reward prediction errors in older compared with younger adults. *The Journal of Neuroscience*, 33(24), 9905–9912. doi:[10.1523/JNEUROSCI.2942-12.2013](https://doi.org/10.1523/JNEUROSCI.2942-12.2013)
- Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, 17(1), 51–72. doi:[10.1162/0898929052880093](https://doi.org/10.1162/0898929052880093)
- Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*, 306(5703), 1940–1943. doi:[10.1126/science.1102941](https://doi.org/10.1126/science.1102941)
- Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology*, 68, 101–128. doi:[10.1146/annurev-psych-122414-033625](https://doi.org/10.1146/annurev-psych-122414-033625)
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(Supplement_3), 15647–15654. doi:[10.1073/pnas.1014269108](https://doi.org/10.1073/pnas.1014269108)
- Grieve, S. M., Williams, L. M., Paul, R. H., Clark, C. R., & Gordon, E. (2007). Cognitive aging, executive function, and fractional anisotropy: A diffusion tensor MR imaging study. *American Journal of Neuroradiology*, 28(2), 226–235. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17296985
- Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35(1), 4–26. doi:[10.1038/npp.2009.129](https://doi.org/10.1038/npp.2009.129)

- Hämmerer, D., & Eppinger, B. (2012). Dopaminergic and prefrontal contributions to reward-based learning and outcome monitoring during child development and aging. *Developmental Psychology*, 48(3), 862–874. doi:10.1037/a0027342
- Hämmerer, D., Li, S.-C., Müller, V., & Lindenberger, U. (2011). Life span differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *Journal of Cognitive Neuroscience*, 23(3), 579–592. doi:10.1162/jocn.2010.21475
- Haruno, M., & Kawato, M. (2006). Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: FMRI examination in stimulus-action-reward association learning. *Neural Networks*, 19(8), 1242–1254. doi:10.1016/j.neunet.2006.06.007
- Head, D., Kennedy, K. M., Rodrigue, K. M., & Raz, N. (2009). Age differences in perseveration: Cognitive and neuroanatomical mediators of performance on the Wisconsin card sorting test. *Neuropsychologia*, 47(4), 1200–1203. doi:10.1016/J.NEUROPSYCHOLOGIA.2009.01.003
- Karrer, T. M., Josef, A. K., Mata, R., Morris, E. D., & Samanez-Larkin, G. R. (2017). Reduced dopamine receptors and transporters but not synthesis capacity in normal aging adults: A meta-analysis. *Neurobiology of Aging*, 57, 36–46. doi:10.1016/j.neurobiolaging.2017.05.006
- Ligneul, R. (2019). Sequential exploration in the Iowa gambling task: Validation of a new computational model in a large dataset of young and old healthy participants. *PLoS Computational Biology*, 15(6), e1006989. doi:10.1371/journal.pcbi.1006989
- Logan, J. M., & Balota, D. A. (2008). Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, and Cognition*, 15(3), 257–280. doi:10.1080/13825580701322171
- Ludvig, E. A., Madan, C. R., McMillan, N., Xu, Y., & Spetch, M. L. (2018). Living near the edge: How extreme outcomes and their neighbors drive risky choice. *Journal of Experimental Psychology: General*, 147(12), 1905–1918. doi:10.1037/xge0000414
- Madden, D. J., Costello, M. C., Dennis, N. A., Davis, S. W., Shepler, A. M., Spaniol, J., ... Cabeza, R. (2010). Adult age differences in functional connectivity during executive control. *Neuroimage*, 52(2), 643–657. doi:10.1016/j.neuroimage.2010.04.249
- Maddox, G. B., Balota, D. A., Coane, J. H., & Duchek, J. M. (2011). The role of forgetting rate in producing a benefit of expanded over equal spaced retrieval in young and older adults. *Psychology and Aging*, 26(3), 661–670. doi:10.1037/a0022942
- Mell, T., Heekeren, H. R., Marschner, A., Wartenburger, I., Villringer, A., & Reischies, F. M. (2005). Effect of aging on stimulus-reward association learning. *Neuropsychologia*, 43(4), 554–563. doi:10.1016/j.neuropsychologia.2004.07.010
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. doi:10.1037/h0043158
- Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the Basal Ganglia. *Annual Review of Neuroscience*, 25, 563–593. doi:10.1146/annurev.neuro.25.112701.142937
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042–1045. doi:10.1038/nature05051
- Ponds, R. W. H. M., Verhey, F. R. J., Rozendaal, N., Jolles, J., & Deelman, B. G. (1992). Brief cognitive screening tests for dementia: Comparison of the mini-mental state examination and the cognitive screening tests. In H. Bouma & J. A. M. Graafmans (Eds.), *Gerontechnology* (pp. 261–264). Retrieved from <http://www.narcis.nl/publication/RecordID/oa: dare:8410>
- Radulescu, A., Daniel, R., & Niv, Y. (2016). The effects of aging on the interaction between reinforcement learning and attention. *Psychology and Aging*, 31(7), 747–757. doi:10.1037/pag0000112
- Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., ... Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex*, 15(11), 1676–1689. Retrieved from <https://doi.org/10.1093/cercor/bhi044>
- Raz, N., Rodrigue, K. M., Kennedy, K. M., Head, D., Gunning-Dixon, F., & Acker, J. D. (2003). Differential aging of the human striatum: Longitudinal evidence. 24(9), 1849–1856. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14561615>

- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts. Retrieved from <http://www.ualberta.ca/~egray/teaching/Rescorla&Wagner1972.pdf>
- Ridderinkhof, K. R., Span, M. M., & van der Molen, M. W. (2002). Perseverative behavior and adaptive control in older adults: Performance monitoring, rule induction, and set shifting. *Brain and Cognition*, 49(3), 382–401. Retrieved from <https://doi.org/10.1006/brcg.2001.1506>
- Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(9), 1430–1443. doi:10.1037/xlm0000518
- Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *Journal of Neuroscience*, 29(48), 15104–15114. doi:10.1523/JNEUROSCI.3524-09.2009
- Rypma, B., Prabhakaran, V., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Load-dependent roles of frontal brain regions in the maintenance of working memory. *NeuroImage*, 9(2), 216–226. doi:10.1006/nimg.1998.0404
- Salat, D. H., Greve, D. N., Pacheco, J. L., Quinn, B. T., Helmer, K. G., Buckner, R. L., & Fischl, B. (2009). Regional white matter volume differences in nondemented aging and Alzheimer's disease. *NeuroImage*, 44(4), 1247–1258. Retrieved from <https://doi.org/10.1016/j.neuroimage.2008.10.030>
- Schmand, B. A., Bakker, D., Saan, R. J., & Louman, J. (1991). De Nederlandse leestest voor Volwassenen: Een maat voor het premorbidie intelligentieniveau/The Dutch adult reading test: A measure of premorbid intelligence. *Tijdschrift Voor Gerontologie En Geriatrie*, 22(1), 15–19.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599. doi:10.1126/science.275.5306.1593
- Schultz, W. (2013). Updating dopamine reward signals. *Current Opinion in Neurobiology*, 23(2), 229–238. doi:10.1016/j.conb.2012.11.012
- Seger, C. A., & Cincotta, C. M. (2006). Dynamics of frontal, striatal, and hippocampal systems during rule learning. *Cerebral Cortex*, 16(11), 1546–1555. doi:10.1093/cercor/bhj092
- Simon, H. A. (1974). How big is a chunk? By combining data from several experiments, a basic human memory unit can be identified and measured. *Science*, 183(4124), 482–488. doi:10.1126/science.183.4124.482
- Simon, J. R., Howard, J. H., & Howard, D. V. (2010). Adult age differences in learning from positive and negative probabilistic feedback. *Neuropsychology*, 24(4), 534–541. doi:10.1037/a0018652
- Simone, P. M., Bell, M. C., & Cepeda, N. J. (2013). Diminished but not forgotten: Effects of aging on magnitude of spacing effect benefits. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 68(5), 674–680. doi:10.1093/geronb/gbs096
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology*, 6(2), 174–215. <http://dx.doi.org/10.1037/0033-2909.87.2.245>
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245–251. Retrieved from <http://dx.doi.org/10.1037/0033-2909.87.2.245>
- Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. xavierdupre.fr. Cambridge, MA: MIT Press. Retrieved from http://www.xavierdupre.fr/enseignement/projet_data/apprentissage_reinforcement_RL-3.pdf
- Tisserand, D. J., van Boxtel, M. P. J., Pruessner, J. C., Hofman, P., Evans, A. C., & Jolles, J. (2004). A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. *Cerebral Cortex*, 14(9), 966–973. doi:10.1093/cercor/bhh057
- Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. *Psychology of Learning and Motivation*, 60, 113–189. doi:10.1016/B978-0-12-800090-8.00004-4
- Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences*, 31(6), 837–862. doi:10.1016/S0191-8869(00)00183-5

- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28(2), 127–154. doi:[10.1016/0749-596X\(89\)90040-5](https://doi.org/10.1016/0749-596X(89)90040-5)
- Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, 71, 1–26. doi:[10.1016/J.COGPSPYCH.2014.01.003](https://doi.org/10.1016/J.COGPSPYCH.2014.01.003)
- van de Vijver, I., Cohen, M. X., & Ridderinkhof, K. R. (2014). Aging affects medial but not anterior frontal learning-related theta oscillations. *Neurobiology of Aging*, 35(3), 692–704 <https://doi.org/10.1016/j.neurobiolaging.2013.09.006>.
- van de Vijver, I., Ridderinkhof, K. R., Harsay, H., Reneman, L., Cavanagh, J. F., Buitenweg, J. I. V., & Cohen, M. X. (2016). Frontostriatal anatomical connections predict age- and difficulty-related differences in reinforcement learning. *Neurobiology of Aging*, 46, 1–12. doi:[10.1016/j.neurobiolaging.2016.06.002](https://doi.org/10.1016/j.neurobiolaging.2016.06.002)
- van de Vijver, I., Ridderinkhof, K. R., & de Wit, S. (2015). Age-related changes in deterministic learning from valenced performance feedback. *Aging, Neuropsychology, and Cognition*, 22(5), 595–619. <https://doi.org/10.1080/13825585.2015.1020917> .
- Wechsler, D. (2008). *Wechsler adult intelligence scale* (4th ed.). San Antonio, TX: Pearson Assessment.
- Weiler, J. A., Bellebaum, C., & Daum, I. (2008). Aging affects acquisition and reversal of reward-based associative learning. *Learning & Memory*, 15(4), 190–197. doi:[10.1111/j.1460-9568.2012.08017.x](https://doi.org/10.1111/j.1460-9568.2012.08017.x)
- Wimmer, G. E., Daw, N. D., & Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. *European Journal of Neuroscience*, 35(7), 1092–1104. doi:[10.1111/j.1460-9568.2012.08017.x](https://doi.org/10.1111/j.1460-9568.2012.08017.x)
- Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*, 338(6104), 270–273. doi:[10.1126/science.1223252](https://doi.org/10.1126/science.1223252)
- Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H.-J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-Dependent long-term memory formation. *Neuron*, 45(3), 459–467. doi:[10.1016/J.NEURON.2005.01.010](https://doi.org/10.1016/J.NEURON.2005.01.010)
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49. doi:[10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)
- Ziegler, D. A., Pigué, O., Salat, D. H., Prince, K., Connally, E., & Corkin, S. (2010). Cognition in healthy aging is related to regional white matter integrity, but not cortical thickness. *Neurobiology of Aging*, 31(11), 1912–1926. Retrieved from <https://doi.org/10.1016/j.neurobiolaging.2008.10.015>