The effects of aging on the interaction between reinforcement learning and attention

Angela Radulescu*, Reka Daniel and Yael Niv

Department of Psychology and Princeton Neuroscience Institute

Princeton University, Princeton NJ, USA

*Correspondence to: angelar@princeton.edu

Abstract

Reinforcement learning (RL) in complex environments relies on selective attention to uncover those aspects of the environment that are most predictive of reward. While previous work has focused on age-related changes in RL, it is not known whether older adults learn differently from younger adults when selective attention is required. In two experiments, we examined how aging impacts on the interaction between RL and selective attention. Younger and older adults performed a learning task in which only one stimulus dimension was relevant to predicting reward, and within it, one 'target' feature was the most rewarding. Participants had to discover this target feature through trial and error. In Experiment 1 stimuli varied on one or three dimensions and participants received hints that revealed the target feature, the relevant dimension, or gave no information. Group-related differences in accuracy and reaction times differed systematically as a function of the number of dimensions and the type of hint available. In Experiment 2 we used trial-by-trial computational modeling of the learning process to test for age-related differences in learning strategies. Behavior of both young and older adults was explained well by a reinforcement-learning model that uses selective attention to constrain learning. However, the model suggested that older adults restricted their learning to fewer features, employing more focused attention than younger adults. Furthermore, this difference in strategy predicted age-related deficits in accuracy. We discuss these results suggesting that a narrower filter of attention may reflect an adaptation to the reduced capabilities of the reinforcement learning system.

Keywords: aging, reinforcement learning, selective attention

Introduction

Studies suggest that healthy aging impacts on reinforcement learning (RL) – the ability to associate stimuli with expected future rewards (Mell et al., 2005; Mata, Josef & Samanez-Larkin, 2011; Samanez-Larkin & Knutson, 2014). When required to learn stimulus-reward associations from feedback, older adults consistently need more trials to reach the same level of performance as younger adults, and exhibit slower reaction times. Previous work has also emphasized that dopamine neurons – which have been implicated in reinforcement learning (see Niv, 2009 for a review) – are gradually lost over the course of the lifespan (Li, Lindenberger & Bäckman, 2010; Eppinger, Hämmerer & Li, 2011). Drawing on these findings, age-related behavioral differences in RL tasks have been linked to a reduced efficacy in reward prediction-error signaling in the human striatum (Eppinger, Schuck, Nystrom & Cohen, 2013). Using a pharmacological manipulation, Chowdhury and colleagues further showed that dopaminergic drugs can restore this signal, and boost the performance of older adults to levels comparable to those observed in younger adults (Chowdhury et al., 2013).

But is a deficit in simple stimulus-reward learning really at the heart of the difficulties that older adults show when learning in the real world? One aspect that might give us pause is that typical RL tasks use simple stimuli that do not reflect the complex nature of our day-to-day environment. Outside the laboratory, a stimulus may not be an isolated tone or a colored shape on a screen. Instead, task-relevant stimuli—for example, the 'dial number' option on a new mobile phone—are embedded within a cluttered environment, among many other stimuli that are not relevant for the task at hand. Moreover, stimuli are multidimensional, with only some dimensions, such as the location of the 'dial' button (but not its font or color) being critical for obtaining reward. In order to learn efficiently, one needs to discover the relevant dimensions via trial-and-error and restrict learning to these, while ignoring other distractors. We have previously

suggested that this process of 'representation learning' depends on the interaction between RL and selective attention. Notably, attention filters what we learn about, and this attention filter is itself dynamically modulated by reinforcement (Wilson & Niv 2012; Niv et al., 2015; Geana and Niv 2015). Here we ask whether older adults differ from younger adults in their selective attention strategies and/or in the efficacy of their ability to learn from feedback in a multidimensional environment in which only some dimensions are relevant for the task at hand.

Several lines of research suggest that the interplay between RL and attention may change with age. Behaviorally, older adults exhibit lower performance on tasks that require internally generating and maintaining task-relevant information (Braver & Barch, 2002; Hampshire, Gruszka, Fallon & Owen, 2008), as well as suppressing task-irrelevant distractors (Gazzaley, Cooney, Rissman & D'Esposito, 2005; Campbell, Grady, Ng & Hasher, 2012; Schmitz, Cheng & DeRosa, 2010). A recent review summarized evidence that older adults compensate for these lapses in cognitive control by relying more on the external environment to provide task-appropriate representations (Lindenberger & Mayr, 2014). At the neural level, it has been suggested that changes in the interaction between DA and the prefrontal cortex (PFC) can account for observed differences in attentional modulation and inhibition of irrelevant stimuli (Li et al. 2010; Cabeza & Dennis 2013; Braver & Barch, 2002). Taken together, these findings suggest that age may strongly affect the interaction between reward learning and attention.

To test this, we compared the behavior of younger and older adults on variants of a probabilistic learning task with multidimensional stimuli. In the first experiment, we varied (1) the number of dimensions that stimuli differ on, and (2) the availability of hints that help curtail learning demands by focusing the subjects' attention on relevant aspects of the stimuli. In the second experiment, we used computational modeling to characterize and compare different

learning strategies in younger and older adults in the most demanding case, in which the relevant dimension is unknown.

General methods

Task

On each trial of the task, participants were presented with three visual stimuli. Stimuli differed along either one or three dimensions (color, shape and texture, Fig. 1). Within each dimension, a given stimulus could have one of three features (e.g. red, green and yellow, Fig. 1). On each trial, participants chose between stimuli that consisted of random combinations of features (e.g. red square with polka dots). Importantly, at any time point, only one dimension of the stimuli determined reward. Specifically, one "target" feature within this “relevant” dimension was more rewarding than the others: choosing the stimulus that contained the target feature led to 75% chance of receiving 1 point (and 0 points otherwise), while choosing either of the other two stimuli was rewarded by 1 point with only 25% chance. Participants were fully informed of these reward probabilities, and the existence of a relevant dimension and target stimulus within it. To maximize the number of points earned, participants had to learn the identity of the target feature and use it to select the correct stimulus on each trial. Participants were asked to make their choice within 2 seconds, after which the trial timed out and the next trial began. To acquire repeated measurements of learning within each participant, we divided the task into several “games”. The identity of the target feature stayed constant throughout a game. Once the game ended, participants were allowed a short, self-paced break and were notified that the relevant dimension and target feature would now be changing. This task is related to the Wisconsin Card Sorting Task that has previously been used to study cognitive flexibility in older adults (Fristoe et al., 1997; Rhodes 2004), with the key difference being that rewards were probabilistic much

like in the weather prediction task (Ashby & Maddox, 2005). The design prolonged the learning process such as to allow the use of computational modeling to analyze the dynamics of learning.

Exclusion criteria

We considered a game to be "learned" if the participant chose the stimulus containing the target feature in each of the last six trials of the game. In both experiments, we excluded participants who learned fewer than 20% of all games, missed more than 10% of the trials, or performed at chance in any of the tasks. Chance was defined as less than 38% accuracy (two standard deviations above the mean of a binomial distribution with $p = 1/3$ and $N = 1000$ trials, matching the average number of trials performed by participants).

Statistics

To quantify effect sizes, unless otherwise noted, we report the following: (1) Hedges' g for independent-sample t-tests, a measure more robust to small samples (Hedges, 1981; Hentschke & Stüttgen, 2011), (2) partial η^2 for ANOVAs, (3) Pearson correlation coefficients, (4) standardized regression coefficients. All data plots show ± 1 SEM in red and 95% confidence intervals in blue.

Experiment 1

In the first experiment, our aim was to separately assess the contributions of reinforcement learning and attention to age-differences in performance of trial and error learning in a multidimensional environment. We tested each participant on five versions of the task, manipulating the number of dimensions along which stimuli varied (one or three, henceforth

abbreviated as 1D and 3D) and the availability of hints that could be used to reduce the computational demands of probabilistic learning (Fig. 1, see also Fig. S1 available online). Throughout, reward contingencies and motor requirements were kept constant.

Participants

33 younger adults (23 female, 10 male; mean age = 23 years; age range = 19 - 37) and 33 older adults (16 female, 17 male; mean age = 69.4 years; age range = 62 - 80) participated in the experiment for either monetary compensation or course credit (younger adults). Older participants were recruited from among members of the Community Auditing Program at Princeton University.

All participants reported normal or corrected-to-normal color vision, were enrolled in an undergraduate program or held at least a university degree, had no history of psychiatric disorders, and provided informed consent. The experiment was approved by the Princeton University Institutional Review Board. The older adult cohort was screened for early-onset dementia using a shortened version of Raven's Progressive Matrices (Raven & Court, 1998). One older adult who scored less than a 5 (out of 18) on this test was excluded from further analysis. Additionally, 2 younger adults and 7 older adults were excluded from further analysis as per the task performance criteria above, yielding a final sample of 31 younger adults and 25 older adults.

Stimuli and procedure

Games in the task were divided into five randomly intermingled conditions with 10 repetitions each.

In the first condition ("feature 1D"), stimuli varied on one dimension (e.g., the 3 distinct shapes; Fig. 1 top). Before the game, a "hint" screen revealed which of the features within this dimension was the target (e.g., 'square'; Fig. 1 and Fig. S1 available online). This condition was thus equivalent to performing a visual search for a predefined feature, and had no learning component: to maximize reward, participants simply had to select the target feature.

In the second condition ("feature 3D"), we again cued participants regarding the target feature, but presented them with three-dimensional stimuli each varying along color, shape and texture (Fig. 1 bottom). Comparing group behavior between the 1D scenario above and the 3D case allowed us to ask whether older adults show a disadvantage when distractor dimensions are present even when no learning is required.

In the third condition ("dimension 1D"), stimuli varied on a single dimension, but instead of being told the identity of the target feature, participants had to learn it from trial and error. This condition is equivalent to a 3-armed bandit task akin to the kinds of tasks that have previously been studied in older adults to characterize deficits in reinforcement learning (Mell et al., 2005; Chowdhury et al., 2013).

The fourth condition ("dimension 3D") involved three-dimensional stimuli and a hint disclosing which dimension is relevant for predicting reward. Participants were thus required to learn the identity of the target feature within that dimension, as in the dimension 1D task. However, because distractor dimensions were present (Fig. 1, bottom), this task required sustained attention to one dimension – to do well, participants had to restrict learning to the cued dimension and ignore the other distracting dimensions.

Finally, in the fifth condition ("full 3D"), participants were presented with three-dimensional cues, and received no information as to the relevant dimension or target feature. This condition is identical to the "dimensions task" we have previously used to investigate the

interaction between selective attention and reinforcement learning in younger adults (Wilson & Niv, 2012; Niv et al., 2015; Geana & Niv, 2015).

Prior to the experimental session, participants were given a tutorial that described the reward structure and informed them about the different conditions. Following the tutorial, participants completed several sample games. They were then tested on 50 games of the task with condition randomized such that within each block of five games, each condition appeared once for a total of 10 games per condition. Each game consisted of a minimum of 8 and a maximum of 25 trials. We defined a correct trial as one in which the participant chose the stimulus containing the target feature. Once a criterion of 8 consecutive correct trials was reached, the game had a 50% chance of ending on any subsequent trial. Games lasted at most 25 trials. The target feature was chosen randomly, avoiding relevant dimension repeats from game to game. Rewards were drawn pseudorandomly such that within each block of eight trials the frequency of presented rewards matched the reward probabilities specified in the design. Once a valid response was registered, stimuli that were not chosen were removed from the screen and after a brief delay the outcome was presented for 0.5 seconds. A new trial started after 0.5 seconds.

Apparatus

Participants sat approximately 50 cm from an LCD monitor and responded on a standard Macintosh keyboard using three adjacent keys corresponding to the left, middle and right stimulus respectively. Stimuli were presented and responses were registered using MATLAB (The MathWorks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Reaction time model

The ex-Gaussian distribution has been proposed as an analysis tool for reaction time that disentangles variance arising from two separate cognitive processes: the *transduction* component, indexed by a positive shift in the Gaussian mean, can be thought of as the time it takes to process the sensory information plus the time required to physically make the response once a decision has been made. The *decision* component, reflected in the exponential skew, is a proxy for the time it takes to represent the task and decide which response to make (Luce, 1986; Lacouture & Cousineau, 2008).

We used maximum likelihood estimation to fit individual reaction time distributions separately within each condition (Fig. S2 available online). By analyzing reaction time data in this way, we sought to dissociate age effects in how the task is represented from perceptual and motor differences. We hypothesized that as representational demands increase with the number of dimensions, older adults would be selectively impaired in the decision component of reaction time.

Results

We first examined the effect of condition on overall accuracy. A 2 (age-group) x 5 (condition) mixed-effects ANOVA (Fig. 2A) revealed a main effect of age ($F(1, 54) = 6.80$, $p = .01$, $\eta_p^2 = .11$), a main effect of condition ($F(4, 216) = 705.66$, $p < .001$, $\eta_p^2 = .93$) and a significant interaction between age group and learning condition ($F(4, 216) = 5.50$, $p < .001$, $\eta_p^2 = .09$).

To better delineate the various determinants of accuracy differences between older and younger adults, we next performed 2 (age-group) x 2 (condition) ANOVAs examining the effect of each progression in task difficulty. We first compared the two “feature” conditions, in which

trial-and-error learning was not necessary since the target feature was disclosed via the hint. In the "feature 1D" condition, both groups performed at ceiling. While we did find a main effect of condition when adding the distractor dimensions in the "feature 3D" case ($M_{feature1D}$ = .993, $M_{feature3D}$ = .987, F(1, 54) = 14.122, p < .001, η_p^2=.21), there was no main effect of group (F(1, 54) = 2.38, p = .13, η_p^2= .04) and no interaction between age and condition (F(1, 54) = .28, p = .60, η_p^2= .01). These results suggest that the presence of extra-dimensional distractors did not specifically impair accuracy in older adults, and also confirms that participants understood the instructions and used the hint correctly.

We next focused on performance differences when participants had to learn the target feature from trial and error, but were cued as to the relevant dimension. In the 1D case (Fig. S1 available online), this amounts to a simple 3-way choice task with binary rewards and fixed reward probabilities. As expected from previous work showing age-related impairments in probabilistic learning, we observed a significant group effect on overall accuracy when comparing the "feature 1D" to the "dimension 1D" condition (main effect of group: F(1, 54) = 5.56, p = .02, η_p^2= .09; interaction: F(1, 54) = 7.74, p = .01, η_p^2= .13).

In the 3D case, the dimension hint helps participants assign credit for a reward to only one of the 3 features of a stimulus: if the reward-relevant dimension is color, feedback for, say, choosing a green square with polka dots, can be correctly assigned to the color 'green', while ignoring the 'square' and 'polka dot' features that act as distractors. Thus the "dimension-3D" case is similar to a 3-way choice task, only with known distractors. As expected, here too we observed a significant group effect on accuracy when comparing the "feature 3D" to the "dimension 3D" condition (main group effect: F(1, 54) = 7.18, p = .01, η_p^2= .12; interaction: F(1, 54) = 8.77, p = .01, η_p^2= .14). These age-related deficits in learning from trial and error

have previously been attributed to a reduced efficacy of dopamine-dependent prediction error signals (Eppinger et al. 2013).

Finally, we compared the "dimension 3D" condition with the "full 3D" condition. In the latter, participants were also required to identify the relevant dimension from trial and error. As a result of this extra demand, we expected the performance of older adults to drop more precipitously than that of younger adults, as compared to the "dimension 3D" case. In line with this prediction, our analysis revealed a main effect of group ($F(1, 54) = 5.73$, $p = .02$, $\eta_p^2 = .10$) and of condition ($F(1, 54) = 183.93$, $p < .001$, $\eta_p^2 = .77$). We also observed a significant interaction between group and condition ($F(1, 54) = 5.89$, $p = .02$, $\eta_p^2 = .10$). Surprisingly however, this interaction was in the direction opposite from what we predicted. That is, when the dimension hint was removed, older adults incurred a smaller additional cost in accuracy than younger adults, performing as well as younger adults on the task ($M_{older} = .48$, $SD_{older} = .09$; $M_{younger} = .49$, $SD_{younger} = .06$, $t(54) = .65$; $p = .52$, $g = .17$). To rule out the possibility of either a floor or ceiling effect driving this interaction, we performed a within-group median split on accuracy (Figure S3 available online). Both older and younger adults were distributed symmetrically around their respective group mean, giving no indication that the observed result was due to boundary effects.

We next analyzed response times using the ex-Gaussian distribution. Examining reaction times revealed more subtle effects of condition than were apparent in overall accuracy. We submitted the fitted skew parameters indexing the decision component of the reaction time to a 2 (age-group) x 5 (condition) mixed-effects ANOVA (Fig. 2B). Paralleling the accuracy results, this initial test revealed a main effect of age ($F(1, 54) = 17.53$, $p < .001$, $\eta_p^2 = .25$), a main effect of condition ($F(4, 216) = 108.15$, $p < .001$, $\eta_p^2 = .67$) and a significant interaction between age group and learning condition ($F(4, 216) = 10.68$, $p < .001$, $\eta_p^2 = .17$). An independent-samples t-

test indicated no group difference between older (M = .13, SD = .05) and younger (M = .15, SD = .06) adults for the decision component in the "feature 1D" condition, t(54) = .81, p = .42, g = .22, indicating that by analyzing reaction time using the ex-Gaussian distribution, we were successful in removing the variance associated with the cost of visual search and response mapping. (Nevertheless, all interaction results reported here for the decision component also held when using raw reaction time as the dependent variable; Fig. S4 available online).

We then performed 2 (age-group) x 2 (condition) ANOVAs paralleling those we reported above for accuracy, to separately assess the effect of each manipulation on the decision component of the reaction time. While in both the "feature 1D" and "feature 3D" conditions older and younger adults performed at ceiling as reflected by average accuracy, we did observe a modest group by condition interaction in the decision component of reaction time ($F(1, 54) = 5.12$, $p = .03$, $\eta_p^2 = .09$), suggesting that older adults required more time to decide on their choice when the stimulus consisted of multiple features.

As expected, we also found a group by condition interaction when comparing both the "feature 1D" and the "dimension 1D" conditions ($F(1, 54) = 53.89$, $p < .001$, $\eta_p^2 = .50$), and when comparing the "feature 3D" condition with the "dimension 3D" condition ($F(1, 54) = 7.2$, $p = .01$, $\eta_p^2 = .18$). These results mirror the accuracy effects, and suggest that having to learn the target feature from feedback impacted older adults' decision time significantly more than it did younger adults'.

Finally, we did not observe an interaction between group and condition when comparing the "dimension 3D" with the "full 3D" condition ($F(1, 54) = .05$, $p = .82$, $\eta_p^2 = .00$). This finding suggests that in the full dimensions task, older adults respond slower than younger adults, but not more so than in a simple probabilistic learning setting in which the relevant dimension is known.

Together, the results of Experiment 1 suggest that while older adults are significantly more impaired than younger adults in simple trial-and-error learning, they *do not* show an additional impairment when required to learn which dimension is relevant to determining reward. In this latter case, their accuracy was not significantly different than that of younger adults, and while they did take significantly longer to respond than younger adults when the relevant dimension was unknown, this group difference was not greater than in the cued dimension case. These results are in line with previous reports of age-related deficits in reinforcement learning (Mell et al., 2005; Eppinger, Schuck, Nystrom & Cohen, 2013) and reveal that, contrary to expectations, attentional demands do not confer differential additional hardship on older adults. As previous work has suggested that attention processes do change during healthy aging, one possibility is that older adults adapt their strategies such as to allow them to perform the full 3D representation learning task better than would otherwise be expected.

Experiment 2

To more precisely understand what strategies may help older adults adjust to multidimensional settings despite a deficit in trial-and-error learning, in a second experiment we focused on the "full 3D" scenario and compared strategies between the groups by fitting RL models to choice data. Testing participants only on this task allowed us to collect more games, and thus identify model parameters with higher precision.

Participants

28 younger adults (17 female, 11 male; mean age = 23.9 years; age range = 20 - 31) and 30 older adults (10 female, 20 male; mean age = 70.1 years; age range = 65 - 80) participated in

the second experiment. All participants reported normal or corrected-to-normal color vision, were enrolled in university or held at least a university degree, had no history of psychiatric disorders, and provided informed consent. The protocol was approved by the Princeton University Institutional Review Board.

In addition to completing the main task, participants in both groups also completed several psychometric tests and questionnaires: (1) A computerized version of the Digit-Symbol substitution test (Salthouse, 1992), (2) a 2-back task (Nystrom, Braver, Sabb, Delgado & Cohen, 2000), (3) the Spot-the-Word test (Baddeley, Emslie & Nimmo-Smith, 1992), (4) the BIS-BAS questionnaire (Carver & White, 1994), (5) a shortened version Raven's Progressive Matrices. As in the first experiment, the older adult cohort was screened for early onset dementia using the Raven's Progressive Matrices. Exclusion criteria were identical to those of Experiment 1. One younger adult and 3 older adults were excluded from the analysis, yielding a final sample of 27 younger adults and 27 older adults.

Stimuli and procedure

Stimuli were identical to the "full 3D" condition in Experiment 1. On each trial, participants were presented with multidimensional stimuli varying on color, shape and texture. One of the three dimensions was used to determine rewards. Within this reward-relevant dimension one target feature had a 75% probability of reward, while all other features had a 25% probability of reward. Participants received no information about the identity of the relevant dimension. Each participant played, on average, 35-50 games for a total of approximately 1,500 trials per participant. The length of a game was fixed at 30, excluding missed trials. The total duration of the experiment was capped at 40 minutes. Once a valid response was registered, the

stimuli that were not chosen were removed from the screen and the outcome was immediately presented for 0.5 seconds. A new trial started after 0.3 seconds.

Apparatus

The apparatus was the same as in Experiment 1.

Model-based analysis

We have previously shown that in this task, RL models that allow for effects of selective attention explain subjects' behavior better than either a naïve RL model that learns values for each of the 27 possible stimuli or a Bayesian ideal-observer model that makes statistically optimal use of information (Wilson & Niv, 2012; Niv et al., 2015; Geana & Niv, 2015). Here we were interested in testing for an effect of age on the width of the 'attentional filter'.

Towards this end, we first compared between two RL models, a "feature RL" (fRL) model that attends uniformly to all three dimensions, and a "feature RL with decay" (fRL+decay) model that emulates selective attention to dimensions that include consistently chosen features (see below). Both models track a weight W for each feature f and calculate the value $V(S)$ of stimulus S as the sum of the weights of its three features $W(f)$, where each stimulus has one feature per dimension. In the fRL model, on every trial, once the outcome for the chosen stimulus is displayed, the weights corresponding to the three features of the chosen stimulus are updated according to

$$W^{new}(f) = W^{old}(f) + \eta(R - V(S)) \quad \forall f \in S_{chosen} \qquad (1)$$

where η is the step size or learning rate parameter and R is the reward (1 or 0 points) on the current trial.

The fRL+decay model is identical to the fRL model, except that it also decays to zero the weights of features that do not appear in the chosen stimulus:

$$W^{new}(f) = (1 - d)W^{old}(f) \quad \forall f \notin S_{chosen} \qquad (2)$$

where d is the decay rate. For both models, at the beginning of each game, the weights are initialized at zero. On each trial, the probability of selecting each of the three available stimuli is calculated using a softmax distribution

$$p(S_{chosen} = S_i) = \frac{e^{\beta V(S_i)}}{\sum_{j=1}^{3} e^{\beta V(S_j)}} \qquad (\varepsilon)$$

where the inverse temperature parameter β captures the noise in the subjects' choices. Thus the fRL model has two free parameters, $\theta_{fRL} = \{\eta, \beta\}$ and the fRL+decay model has three free parameters, $\theta_{fRL+decay} = \{\eta, \beta, d\}$.

Importantly, the decay rate d dictates the width of an implicit attentional filter (Fig. S5 available online). To understand this mechanism, it is instructive to consider two hypothetical consecutive trials, t and $(t+1)$, in which a participant might choose stimuli such that only one feature – for example red – appears in the chosen stimulus on both trials. A decay rate of zero reduces the fRL+decay model to simple fRL and means that although the two features that co-

occurred with red on trial t did not appear on trial $(t+1)$, their weights remain unchanged on that second trial. At the other extreme, a decay rate of 1 means that on trial $(t+1)$ all weights except those of features of the most recently chosen stimulus are set to 0, in effect erasing the values learned on trial t for features other than 'red'. This is tantamount to a narrow attention filter that, across trials in which 'red' is consistently chosen, only accumulates value for that feature, effectively attending only to color information. Intermediate values of d smoothly interpolate between these two extremes, with higher decay rates corresponding to more 'focused attention' as reflected in high weights for fewer (recently chosen) features. Another way to view our model is that decay emulates attention with a one trial delay; i.e. we use the features the subject chooses on trial $t+1$ to infer what they attended to on trial t, and decay the learning that was done at t to non-attended dimensions. Such model-based inferences are necessary because we do not have direct access to participants' attention (see Niv et al., 2015 for additional discussion).

To compare between models based on their predictive accuracy, we used participants' trial-by-trial choice behavior to fit the parameters that maximize the likelihood of each subject's choices (Daw, 2011). As the models had different numbers of parameters, we compared models using a leave-one-game-out cross-validation approach: for every participant and every game, we fit the model to all data *excluding* that game. The model and its maximum-likelihood parameters were then used to assign likelihood to the trials of the left-out game. We repeated this procedure for each game, divided the resulting total likelihood by the number of trials N to yield the geometric average of the likelihood per trial. This is a quantity that varies between 0 and 1, and roughly corresponds to the average probability with which the model predicted the choices of the participant (1/3 is chance). This quantity was then used for model comparison, with the model that best predicts participants' behavior deemed the winning model. With this model in hand, we

once again applied maximum likelihood parameter estimation, this time using all available data for each participant, to obtain individual parameters for every participant in each group. All optimizations were carried out using MATLAB's *fmincon* function.

Finally, to restrict the fitting as much as possible to trials in which the participants were still learning (rather than simply selecting the target feature), and to be able to compare our results here with those from Experiment 1, model fitting and model comparison analyses were done after imposing a *post hoc* learning criterion of 8 correct trials in a row and capping the length of each game at 25 (as was the case in Experiment 1). All model-based results reported here also hold without this modification.

Results

Independent-sample t-tests showed that the average accuracy of older adults was significantly lower than that of younger adults ($M_{older} = .42$, $SD_{older} = .04$; $M_{younger} = .46$, $SD_{younger} = .04$, $t(52) = 3.49$; $p = .001$, $g = .94$). Older adults also learned fewer games than younger adults ($M_{older} = .39$, $SD_{older} = .09$; $M_{younger} = .49$, $SD_{younger} = .09$; $t(52) = 4.2$, $p < .001$, $g = 1.13$). These results were contrary to our findings in the first experiment, and suggested that failing to detect a difference in accuracy in Experiment 1 might have been due to the smaller number of games (10 per participant compared to 35-40 in the second experiment). Nevertheless, since Experiment 1 established that reward learning is significantly more impaired in older adults, we were interested in analyzing the differential contributions of learning and selective attention to task performance in older adults versus younger adults.

To determine which learning strategy best describes trial-by-trial behavior for each group, we performed a within-group model comparison by taking the average cross-validated likelihood per trial for each model and submitting it to a repeated measures ANOVA (Fig. 3). In

both groups the fRL+decay model predicted participants' choices better than the fRL model (younger adults: $F(1, 1248) = 42.31$, $p < .001$, $\eta_p^2 = .46$, Fig. 3A; older adults: $F(1, 1248) = 23.32$, $p < .001$, $\eta_p^2 = .47$, Fig. 3B). This finding is in line with previous work in which we have shown that strategies that incorporate attentional mechanisms better explain behavior in the task (Wilson & Niv 2012; Niv et al. 2015). Importantly, we observed no significant difference between the groups in the average per-trial likelihood of the fRL+decay model ($F(1, 1248) = .66$, $p = .42$, $\eta_p^2 = .018$). In other words the fRL+decay captured both groups' strategy equally well and could thus be used to assess group differences in fit parameters.

To test for group differences in the breadth of attention for learning, we compared the decay rate parameters for the two groups. A Mann-Whitney test indicated that older adults had significantly higher decay rates than younger adults (older adults median = .52, younger adults median = .42, $U = 192.0$, $p = .002$, $r = .40$, Fig. 4A). This suggests that older adults utilize narrower stimulus representations in trial-and-error learning in multidimensional environments. The results of all group tests on fit parameters are summarized in Table 1.

If the model indeed captures aspects of subjects' strategy that are relevant to behavior, it should be able to reproduce the qualitative patterns observed in the data. To test this, we used individual fit parameters drawn from each group to simulate 54 agents (27 per group), and computed their average learning curves. We found that the model could perform the task at a level comparable to that of participants, slightly undershooting the performance of younger adults, but capturing that of older adults. Importantly, the model reproduced the general performance differences between older and younger adults (Fig. 4B).

To assess the specific effect of the decay rate parameter on task performance in our data, we took a multiple linear regression approach. In particular, each participant's accuracy was regressed on learning rate, decay rate and inverse temperature estimated from the fRL+decay

model. The results of the regression indicated that the three parameters explained 49% of the variance ($R^2 = .49$, $F(3, 50) = 16.0$, $p < .001$). We found that lower decay rates predicted higher accuracies ($\beta = -.52$, $p < .001$), suggesting that even after taking into account possible effects of learning rate and inverse temperature, the width of the attentional filter modulates performance in the task. We also found that learning rate did not significantly predict accuracy ($\beta = .077$, $p = .61$), while inverse temperature did ($\beta = -.53$, $p < .001$). Including both the learning rate and the inverse temperature parameters in the regression was necessary to isolate the effect of decay on performance. However, because these parameters are not completely separable in the model (that is, they cannot be precisely estimated independently of each other; Daw, 2011), we cannot make strong claims about any observed effect of learning rate or inverse temperature alone. Importantly, we repeated the above analysis within each group, and found that decay remained a significant predictor of performance within both younger ($\beta = -.59$, $p < .01$) and older adults ($\beta = -.35$, $p < .01$). Taken together, our results suggest that more focused attention during learning in older adults can, in part, explain the observed decrease in task performance.

Finally, we investigated whether the decay rate reflects a deficit in working memory rather than more focused attention. To test this, we regressed decay rate as estimated from the dimensions task on the working memory score measured using the 2-back task, and included age as a covariate in the regression. We found that age ($\beta = .09$, $p < .05$), but not working memory ($\beta = -.02$, $p = .19$) significantly predicted decay rate, suggesting that age differences in the decay rate parameter are not due to working memory impairments in older adults.

Discussion

We studied how age affects reinforcement learning in multidimensional environments that characterize real-world learning and decision-making scenarios. In Experiment 1, we tested young and older adults on a set of probabilistic learning tasks in which we manipulated the number of stimulus dimensions and the availability of hints about the identity of the target feature or relevant dimension. Our aim was to dissociate the relative contribution of reinforcement learning and representation learning processes to age-related differences in task performance. We found that age differences in both accuracy and reaction time depended on the extent to which reward learning was required to solve the task. Surprisingly, adding representation learning to the demands of the task did not affect the performance of older adults more than it did the performance of younger adults. The results of the first experiment therefore suggested that older adults might adapt to deficits in reinforcement learning such as to reduce the burden on this mechanism in multidimensional environments.

To test this hypothesis, in Experiment 2 we modeled choice data of a new group of participants who performed the full three-dimensional representation learning and reinforcement learning task without any hints regarding the identity of the target feature or relevant dimension. We found that the behavior of both groups was well described by a reinforcement learning model that emulates an attentional filter by decaying the value of unchosen options to zero. Group differences in the decay rate suggested that older adults employ more focused attention—they are more likely to maintain high values for single features rather than combinations of features. Moreover, this difference in strategy came at a cost: more focused attention at least partially explained the lower performance of older adults in our task.

Two mechanistic explanations are consistent with higher decay rates in older adults: participants could employ narrower selective attention at the time of learning, attributing the

reward to fewer stimulus features; or they could be more likely to forget recently learned feature-reward associations. While further work is necessary to precisely distinguish between the two, both lead to a strategy closer to serial hypothesis testing (Wilson & Niv, 2012) in which older adults attend to single features when learning from reinforcement, while younger adults may be learning about multiple chosen features at once. Ignoring the relationship between reward feedback and incidental features of the chosen stimulus (i.e., unattended features that were not the responsible for it being chosen) may be detrimental if participants have learned the wrong task representation. For instance, a participant could be focusing on the red color when trying to maximize reward, and not notice that, in fact, rewards are obtained more often when the red stimulus happens to be a square. Learning about incidental features would enable more efficient switching to other potentially rewarding features. In this sense, narrowly focused attention could pose difficulties for older adults. However, this is not always the case, and in some situations a narrower focus of attention may be normative. While in our specific task such focused attention is not statistically optimal (Niv et al., 2015), our findings are consistent with a recent proposal that in older adults, general models of the world that have been learned over the lifespan reduce the need to rely on sensory updating (Moran, Symmonds & Dolan, 2014). Focusing on fewer aspects of the environment during learning can, in fact, be seen as an adaptation to the structure of real-world tasks, where correct performance might often depend on only few attributes. Mata and colleagues have termed this idea 'ecological rationality.' They make a compelling case for the argument that age-related deficits in strategy use may not necessarily be due to impaired decision making, and that decision strategies can only be evaluated relative to the environment in which they are used (Mata et al., 2012). A striking such example is work by Worthy and Maddox, who show that older adults perform better than younger adults in a task with complex structure that favors a Win-Stay-Lose-Shift strategy, which older adults are more likely to

employ, over a reinforcement learning strategy (Worthy et al., 2011; Worthy & Maddox, 2012). An interesting avenue of future research would thus be to characterize the statistics of natural tasks that older adults have learned to engage in. An ecological rationality view suggests that older adults should be just as good as younger adults at learning new tasks in which previously learned structure could be 'recycled'.

The idea that older adults may display more focused attention in certain situations has been suggested before, in work examining age differences in category learning. For instance, Glass and colleagues argue that when older adults are trained to categorize exemplars from two prototypes, they take into account fewer stimulus dimensions (Glass et al., 2012). Our findings suggest that this strategy also manifests during sequential learning and decision-making, therefore laying the groundwork for a number of future questions: is more focused attention in older adults accompanied by less or more frequent attention switching, as compared to younger adults? And if so, is there an age difference in how much feedback is needed to redirect attention, as suggested by studies that report a tendency to perseverate in older adults (Ridderinkhof, Span & Van Der Molen, 2002; Rhodes, 2004)? In light of our results, perseveration may be attributed to a more rigid focus of attention that prevents the formation of alternative representations, because it filters out incidental learning. While less efficient, this could reflect an adaptation to the reduced efficacy of dopaminergic signaling (Li et al., 2010) in which selective attention is deployed during learning so as to tax mechanisms subserved by dopamine as little as possible.

Finally, the finding that older adults are more likely to filter out information may seem at odds with a broad literature documenting age-related deficits in suppressing task-irrelevant distractors (Gazzaley, Cooney, Rissman & D'Esposito, 2005; Campbell, Grady, Ng & Hasher, 2012; Schmitz, Cheng & DeRosa, 2010). Lindenberger and Mayr have suggested that the

inability of older adults to ignore visual distractors is linked to a broader developmental trend in which older adults shift from internal cognitive control to relying more on the environment to provide appropriate task representations (Lindenberger & Mayr, 2014). However, increased focus and increased distractibility could both result from an attentional system that cannot allocate attention to multiple items, but rather can only maintain narrow, rigid hypotheses. In such conditions, a distractor that captures attention would do so more strongly, and to the exclusion of the task otherwise being performed, leading to apparent distractibility. When attention can be maintained more broadly, to the task-relevant stimuli and also to incidental features, the effect of attention-grabbing distractors is mitigated.

Another way to reconcile the idea of increased environmental reliance with narrower attention when learning task sets concerns the broader question of how tasks are represented in the brain (Wilson et al. 2014). In a recent paper, Mayr and colleagues have suggested an intriguing explanation for age-related increases in task-switching reaction time costs: older adults may not fully represent the relevant task states, opting for a simpler structure at the expense of flexibility (Mayr et al. 2015). Our experiment was explicitly designed to study how participants learn to represent a new task. Narrow attention in our case has the effect of preventing complex stimulus-reward associations from forming (e.g. the participant is more likely to learn that red predicts reward, instead of red and polka dots predict reward). This narrowness limits internal representations, and simplifies the task as much as possible. The trade-off is that such simple representations may not allow for flexibility in learning new tasks that require de-aliasing similar percepts. That is, older adults might have difficulty when choosing the correct action requires learning about a second, disambiguating feature. Finally the process of learning to attend, which we study here, is different from maintaining (instructed) attention in the face of distraction – we suggest that older adults may filter out important relationships between reward feedback and

incidental cues, while at the same time they may erroneously focus attention on incidental distractors.

A growing interest in applying RL methods to the study of cognitive aging has bridged knowledge about dopaminergic loss in older adults and deficits in trial and error learning (Shohamy and Wimmer 2013). Our study focused on how age-related impairments in RL might play out in multidimensional environments where, in addition to trial-and-error learning, one must learn the relevant task representations. In such cases, attentional mechanisms have been hypothesized to interact with RL so as to allow more efficient learning (Wilson & Niv, 2011; Niv et al., 2015; Geana & Niv, 2015). The present work provides evidence that aging is accompanied by a narrowing of attention during reinforcement learning, perhaps in order to adapt to impairments in neural trial-and-error learning mechanisms.

Acknowledgements

This research was supported by a New Scholar in Aging award to Y.N. from the Ellison Medical Foundation. We are grateful to Stephanie Chan, Nicolas Schuck, Amitai Shenhav, Rachel Wu and two anonymous reviewers for very helpful comments on previous versions of the manuscript.

References

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology*, *32*(1), 55-65.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience & Biobehavioral Reviews*, *26*(7), 809-817.

Cabeza, R., & Dennis, N. A. (1994). Frontal lobes and aging : Deterioration and Compensation. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (2nd ed.). Oxford University Press.

Campbell, K. L., Grady, C. L., Ng, C., & Hasher, L. (2012). Age differences in the fronto-parietal cognitive control network: implications for distractibility. *Neuropsychologia*, *50*(9), 2212-2223.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319.

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.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision making, affect, and learning: Attention and performance XXIII*, *23*, 3-38.

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), 1-17.

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.

Fristoe, N. M., Salthouse, T. A., & Woodard, J. L. (1997). Examination of age-related deficits on the Wisconsin Card Sorting Test. *Neuropsychology*, *11*(3), 428.

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298-1300.

Geana, A., & Niv, Y. (2015). Causal Model Comparison Shows That Human Representation Learning Is Not Bayesian. In *Cold Spring Harbor symposia on quantitative biology* (p. 024851). Cold Spring Harbor Laboratory Press.

Glass, B. D., Chotibut, T., Pacheco, J., Schnyer, D. M., & Maddox, W. T. (2012). Normal aging and the dissociable prototype learning systems. *Psychology and Aging*, *27*(1), 120.

Hampshire, A., Gruszka, A., Fallon, S. J., & Owen, A. M. (2008). Inefficiency in self-organized attentional switching in the normal aging population is associated with decreased activity in the ventrolateral prefrontal cortex. *Journal of Cognitive Neuroscience*, *20*(9), 1670-1686.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*(2), 107-128.

Hentschke, H., & Stüttgen, M. C. (2011). Computation of measures of effect size for neuroscience data sets. *European Journal of Neuroscience*, *34*(12), 1887-1894.

Lacouture, Y., & Cousineau, D. (2008). How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times.*Tutorials in Quantitative Methods for Psychology*, *4*(1), 35-45.

Li, S. C., Lindenberger, U., & Bäckman, L. (2010). Dopaminergic modulation of cognition across the life span. *Neuroscience & Biobehavioral Reviews*, *34*(5), 625-630.

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: is there a dark side to environmental support?. *Trends in Cognitive Sciences*, *18*(1), 7-15.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization* (No. 8). Oxford University Press.

Mata, R., Josef, A. K., Samanez-Larkin, G. R., & Hertwig, R. (2011). Age differences in risky choice: a meta-analysis. *Annals of the New York Academy of Sciences*, *1235*(1), 18-29.

Mata, R., Pachur, T., Von Helversen, B., Hertwig, R., Rieskamp, J., & Schooler, L. (2012). Ecological rationality: a framework for understanding and aiding the aging decision maker. *Frontiers in Neuroscience*, *6*, 19.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, *22*(4), 796.

Mayr, U., Spieler, D. H., & Hutcheon, T. G. (2015). When and why do old adults outsource control to the environment?. *Psychology and Aging*, *30*(3), 624.

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.

Moran, R. J., Symmonds, M., Dolan, R. J., & Friston, K. J. (2014). The Brain Ages Optimally to Model Its Environment: Evidence from Sensory Learning over the Adult Lifespan. *PLoS computational biology*, *10*(1), e1003422.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139-154.

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C.

(2015). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. *The Journal of Neuroscience*, *35*(21), 8145-8157.

Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage*, *11*(5), 424-446.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437-442.

Raven, J. C., & John Hugh Court. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford Psychologists Press.

Rhodes, M. G. (2004). Age-related differences in performance on the Wisconsin card sorting test: a meta-analytic review. *Psychology and Aging*, *19*(3), 482.

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.

Rieckmann, A., & Bäckman, L. (2009). Implicit learning in aging: extant patterns and new directions. *Neuropsychology review*, *19*(4), 490-503.

Salthouse, T. A. (1992). What do adult age differences in the Digit Symbol Substitution Test reflect?. *Journal of Gerontology*, *47*(3), P121-P128.

Samanez-Larkin, G. R., & Knutson, B. (2014). Reward processing and risky decision making in the aging brain. *The Neuroscience of Risky Decision Making*.

Schmitz, T. W., Cheng, F. H., & De Rosa, E. (2010). Failing to ignore: paradoxical neural effects of perceptual load on early attentional selection in normal aging. *The Journal of Neuroscience*, *30*(44), 14750-14758.

Schuck, N. W., Frensch, P. A., Schjeide, B. M. M., Schröder, J., Bertram, L., & Li, S. C. (2013).

Effects of aging and dopamine genotypes on the emergence of explicit memory during s equence learning. *Neuropsychologia*, *51*(13), 2757-2769.

Shohamy, D., & Wimmer, G. E. (2013). Dopamine and the cost of aging. *Nature*, *201*, 3.

Wilson, R. C., & Niv, Y. (2012). Inferring relevance in a changing world. *Frontiers in Human Neuroscience*, *5*.

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, *81*(2), 267-279.

Worthy, D. A., Gorlick, M. A., Pacheco, J. L., Schnyer, D. M., & Maddox, W. T. (2011). With age comes wisdom decision making in younger and older adults. *Psychological science*, *22*(11), 1375-1380.

Worthy, D. A., & Maddox, W. T. (2012). Age-based differences in strategy use in choice tasks. *Frontiers in neuroscience*, *5*(145), 1-10.

Group	Learning rate (η)	**Decay rate (d)**	Inverse temperature (β)
YA	.11 ± .01	**.45 ± .017**	12.52 ± .68
OA	.12 ± .01	**.56 ± .03**	14.06 ± 1.65
	p = .49, r = .005	**p = .002, r = .40**	p = .41, r = .03

Table 1. Best-fit values of the free parameters of the fRL+decay model (mean ± SEM) across groups. Bottom: results of the corresponding Mann-Whitney test for group differences for each parameter.

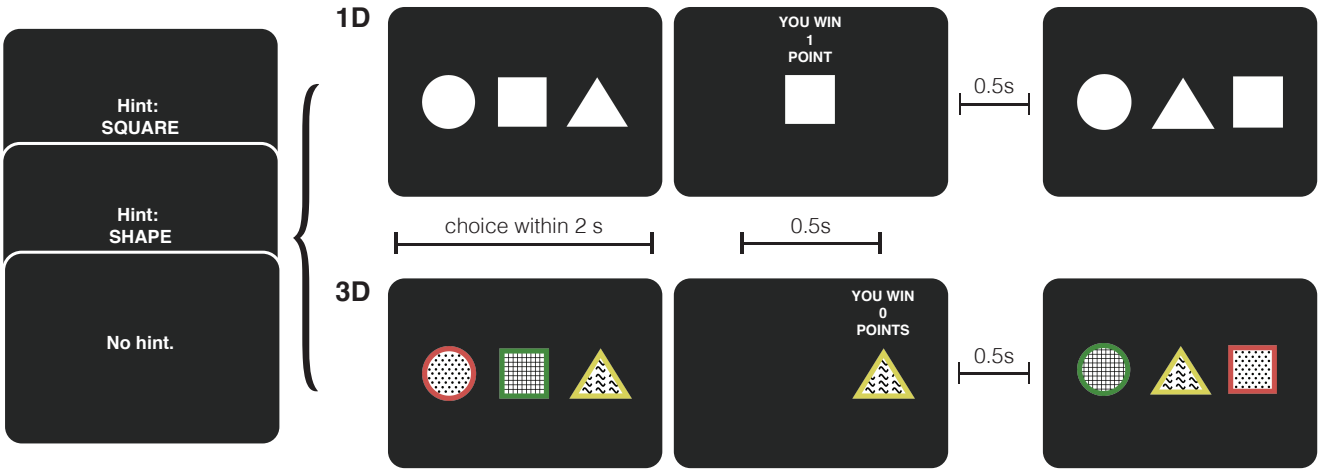

Figure 1. **Outline of the task.** In experiment 1, at the start of each game participants were given a 'hint' regarding the target feature, the dimension of the target feature, or else no hint was given. On each trial, participants chose between three stimuli that varied along a single dimension (e.g., shape) in the 1D case, or along three dimensions (shape, color and texture) in the 3D case. Participants received binary reward feedback, winning either one or zero points on every trial, with reward probability depending on whether they chose the stimulus that contained the target feature. The game ended when the participant reached a performance criterion, or after 25 trials. A new game began with a signaled rule change followed by a new hint screen. Experiment 2 had the same structure, except that all games involved three-dimensional stimuli, no hints, and lasted 30 trials regardless of performance.

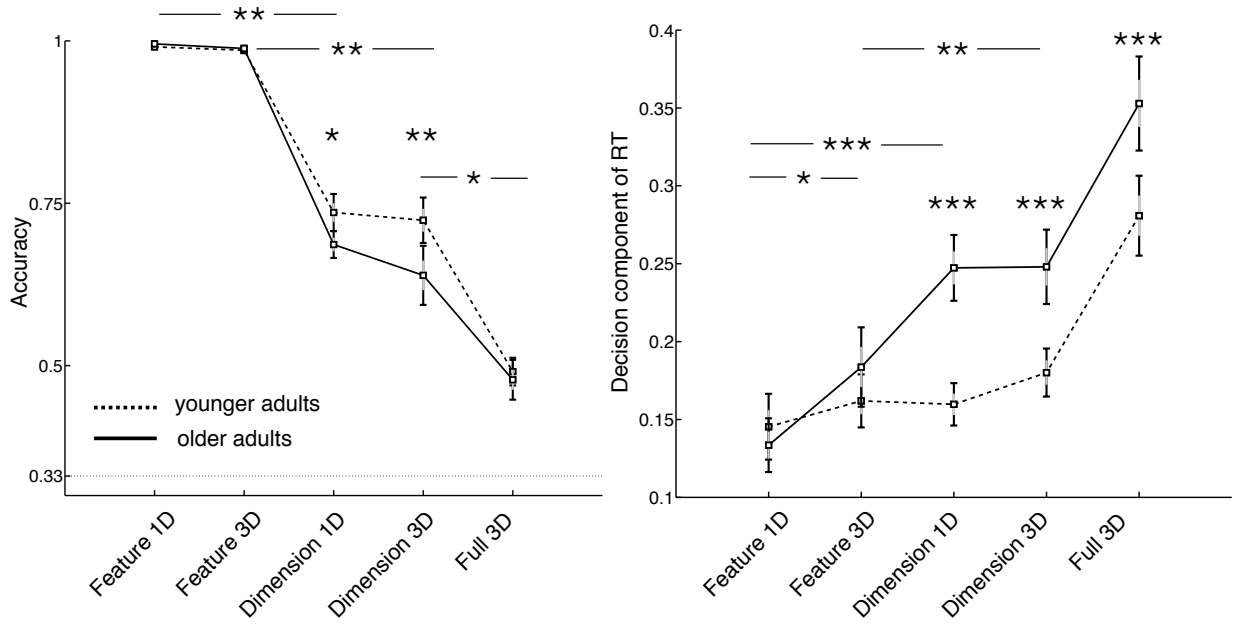

Figure 2. Performance of younger adults (dashed) and older adults (solid) in Experiment 1. (A) Average accuracy in each of the five task conditions. Dotted line indicates chance performance. **(B)** Decision component of reaction time by task condition. Error bars indicate one SEM (gray) and 95% confidence intervals (black). Asterisks indicate significant interactions (lines) and differences between groups within each condition (*p < .05, **p < .01, ***p < .001).

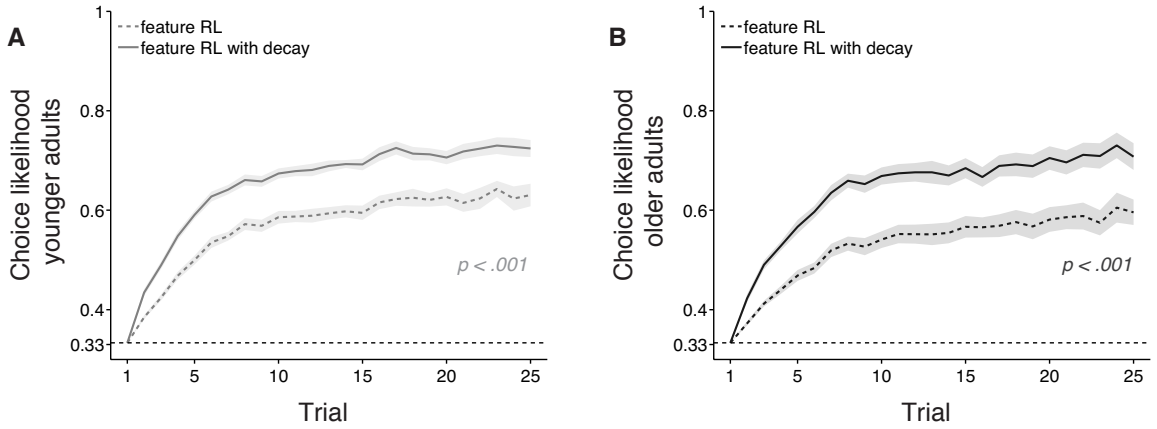

Figure 3. Model comparison. Likelihood per trial as a function of trial within a game for **(A)** Younger adults and **(B)** Older adults. In either group, the data heavily favor the *fRL+decay* model. Dashed line: chance (33%); shading: SEM; p-value corresponds to a repeated measures ANOVA (model x trial).

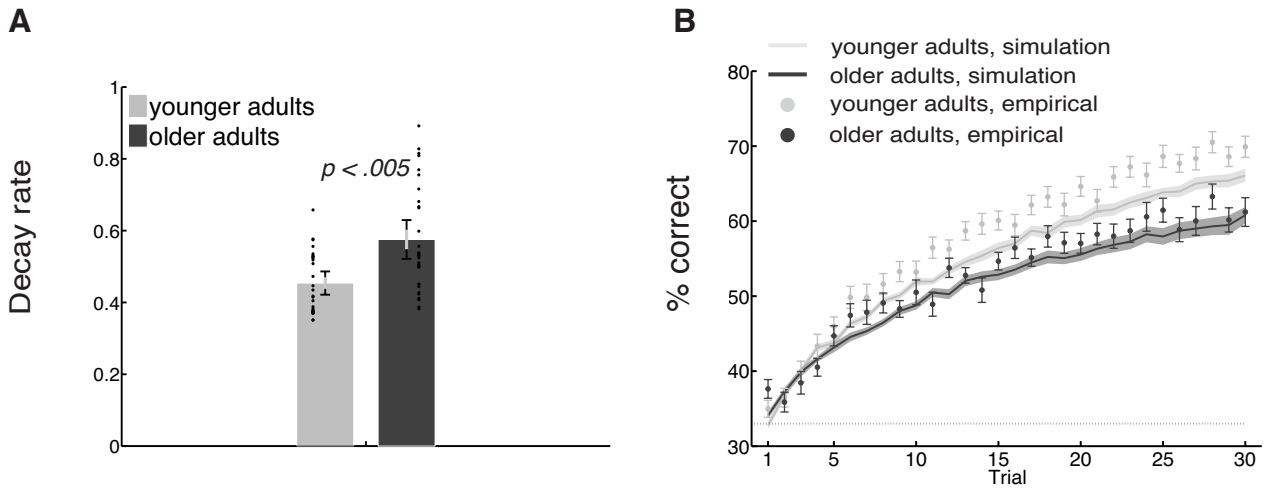

Figure 4. **Age-related differences in decay rate and simulated learning curves. (A)** Group difference in decay rate. Error bars: SEM (gray) and 95% confidence interval (black). Black dots show the decay rate estimates for each participant. **(B)** Average learning curves for participants performing the task (N=27 for each group) and for simulated agents (27 per group) performing 40 games of the task with individual fit parameters within each group. Single data points indicate the empirical performance, and continuous lines indicate simulated performance. Dotted line: chance (33%), error bars and shading: SEM.