Age-Related Differences in Credit Assignment: A pilot study of model-based and model-free learning

Colleen E. Dollst
Institute of Psychology
Universität Hamburg
Hamburg, Germany
colleen.dollst@uni-hamburg.de

Paul Schreiner
Institute of Psychology
Universität Hamburg
Hamburg, Germany
paul.schreiner-2@uni-hamburg.de

Rani Moran
Department of Psychology
Queen Mary University of London
London, UK
r.moran@qmul.ac.uk

Nicolas W. Schuck
Institute of Psychology
Universität Hamburg
Hamburg, Germany
nicolas.schuck@uni-hamburg.de

Xiangjuan Ren
Institute of Psychology
Universität Hamburg
Hamburg, Germany
xiangjuan.ren@uni-hamburg.de

Alexa Ruel
Institute of Psychology
Universität Hamburg
Hamburg, Germany
alexa.ruel@uni-hamburg.de

Abstract

Credit Assignment (CA) – the ability to assign value to the reward-generating aspects of an environment or action – is essential for adaptive decision-making but becomes challenging in complex, multi-step environments. Previous research has found that younger adults can flexibly switch between Model-Free (MF) and Model-Based (MB) strategies to solve the credit assignment problem. In contrast, older adults are less likely to exhibit such flexibility, potentially relying more on MF learning due to task representation difficulties. This pilot study explored age-related differences in CA by adapting a dual-bandit task designed to assess MF and MB contributions and their interactions with behavior. An initial sample of six younger (19-25 years) and six older adults (67-73 years) completed a sequential decision-making task involving binary choices between bandits (or pairs) and inferring a hidden option based on observed outcomes. We could assess the contributions of the MF and MB systems by designing two scenarios based on how the chosen bandit and its counterpart were repeated in successive trials. Preliminary results revealed evident MF contributions in both younger and older adults. However, our conclusion of MB contributions was obscured due to the entanglement of the MF learning. Additionally, we found older adults exhibited slower response times and significant learning across blocks, suggesting greater cognitive effort than younger adults. This study provides initial evidence for potential age-related shifts in CA mechanisms, with significant implications for future research. The need for further work integrating neuroimaging and computational modeling to disentangle MF and MB contributions underlying CA in aging is underscored by our findings.

Keywords: credit-assignment, model-based learning, model-free learning, decision-making, aging

Acknowledgements

This work was supported by a grant from Hamburg University's Ideas and Venture Fund (IVF) for Early Career Researchers, awarded to Dr. Xiangjuan Ren and Dr. Alexa Ruel.

1 Background

A key feature of human intelligence is the ability to learn which actions are most rewarding in environments with different levels of complexity. Identifying which action(s) led to positive or negative outcomes and assigning value to those actions based on their outcome is a process known as *credit assignment (CA)* (Moran et al., 2019; Sutton, 1984). Although straightforward in simple environments, credit assignment becomes harder in complex multi-step scenarios like cooking, navigating an operating system, or when actions and outcomes are separated in time.

Within the *dual reinforcement learning* perspective, two distinct systems have been used to describe human decision-making behavior in these types of complex environments: *Model-Free (MF)* and *Model-Based (MB)* decision-making (Daw et al., 2011; Gershman et al., 2014). During MB decision-making, the agent builds an internal model of the environment and updates action values accordingly, making it adaptable when the environment changes (Daw et al., 2005; Gläscher et al., 2010). This flexibility allows the agent to think ahead (plan) about the future consequences of potential actions. Conversely, a MF decision-making strategy relies solely on trial-and-error choices, making the agent more likely to repeat previously rewarded actions (Daw et al., 2005; Dolan & Dayan, 2013). Previous work examining decision-making across the lifespan reveals that while younger adults flexibly arbitrate between MF and MB strategies, adapting to the demands of the environment, older adults rely predominantly on MF learning (Bolenz et al., 2019; Eppinger et al., 2013). This shift towards simpler strategies has been attributed to older adults' difficulty representing the task environment (Hämmerer et al., 2019; Ruel et al., 2023).

While previous studies have shown that younger adults employ Model-Based (MB) retrospective inference to resolve uncertainty, which subsequently guides Model-Free (MF) credit assignment in a sequential decision-making task (Moran et al., 2019), the interaction between these systems in aging adults remains poorly understood. This study aims to address this gap by focusing on the credit assignment abilities of older adults and further disentangling the distinct contributions of *Model-Free credit assignment (MFCA)* and *Model-Based credit assignment (MBCA)* mechanisms in older and younger adults. The study's unique focus on this interaction could contribute to a deeper understanding of cognitive aging.

2 Method

Based on previous work by Moran and colleagues (2019), 12 pilot participants (six younger and six older adults) completed a modified restless dual-bandit task designed to discriminate between MBCA and MFCA. After learning the deterministic transition structure of the task (Figure 1A) with 100% accuracy, participants completed seven decision-making blocks containing 336 *standard trials* and 168 *uncertain trials* (Figure 1B). Each of the four bandit images in the transition structure (top row of Figure 1A) has two distinct counterpart bandit images, each sharing a different joint outcome image (bottom row of Figure 1A). For example, the bandit image 'bird' is paired with 'tree,' sharing the joint outcome image 'nest,' and its counterpart 'female,' sharing a different joint outcome image, 'airplane.'

During standard trials, either of the two counterpart bandits sharing the same joint outcome image could be displayed. *Standard* and *uncertain trials* were interspersed throughout the blocks, with uncertain trials occurring every third trial. In *standard trials*, after choosing between two bandit images (3s decision time), participants viewed two outcome images associated with their choice successively, the *unique outcome image* specific to the chosen bandit first, followed by the *joint outcome image* shared by both bandit images (see Figure 1B).

The progression of standard trials created two distinct scenarios: 1) the same bandit image chosen in the previous trial and one of its counterpart bandits, which shares a joint outcome image, were displayed in the current trial; 2) a different bandit image, sharing the joint outcome image with the previously chosen bandit, and its counterpart bandit were displayed in the current trial. These scenarios enable the separate evaluation of model-free (MF) and model-based (MB) contributions. Specifically, the main effect of the joint outcome image—whether it was rewarded or not—served as an indicator of the involvement of the two systems (see Results in Figure 3).

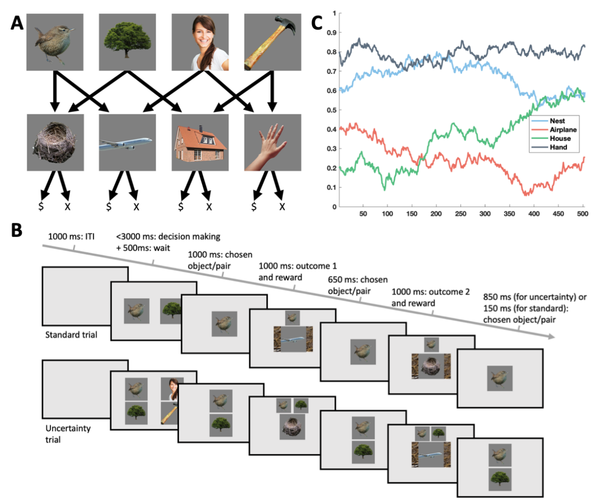

Figure 1: *Task Design and Structure.* ***A.*** *The semantic deterministic transition structure. During the initial learning phase, participants learned the transitions between bandit images (four top-row images) and the four outcome images (bottom-row images). The relationship between the bandit and outcome images was rooted in semantics associations, where the bandit images (e.g., bird) were intuitively linked to their corresponding outcome images (e.g., nest or plane) to help participants learn the transition structure of the task.* ***B.*** *Standard and uncertain trials.* ***C.*** *Reward probabilities of four outcome images. The reward probability for each outcome image drifted according to an independent Gaussian random walk across trials.*

In *uncertainty trials*, participants selected between pairs of bandit images, with the specific bandit image chosen randomly among the pair (the 'glitch'). Contrary to the standard trials, participants first saw the joint outcome image to the selected bandit image pair, followed by the outcome image unique to the chosen bandit image through the 'glitch.' The present analyses focused exclusively on *standard trials*, which is irrelevant to the MB retrospective inference on MFCA from the *uncertainty trials*. Notably, regardless of the trial type, the reward probabilities for the outcome images drifted independently across trials according to Gaussian random walks with reflecting bounds (Figure 1C).

3 Results

Both younger adults' and older adults' accuracies were significantly above chance (Figure 2A; younger adults: 0.75 ± 0.05 (Mean ± SEM), one-sample t-test, $t(5) = 4.65$, $p = 0.006$; older adults: 0.73 ± 0.06, $t(5) = 3.64$, $p = 0.015$). There was no significant difference in accuracy across age groups (independent t-test: $t(10) = 0.26$, $p = 0.80$), but younger adults exhibited significantly faster RTs compared to older adults (Figure 2C; younger adults: $0.81s \pm 0.07s$, older adults: $1.24s, \pm 0.05s$; independent t-test: $t(10) = -4.89$, $p < 0.001$). Furthermore, older adults' accuracy improved significantly from the first to the final block (paired-sample t-test: $t(5) = -3.42$, $p = 0.019$), suggesting a learning effect (Figure 2B).

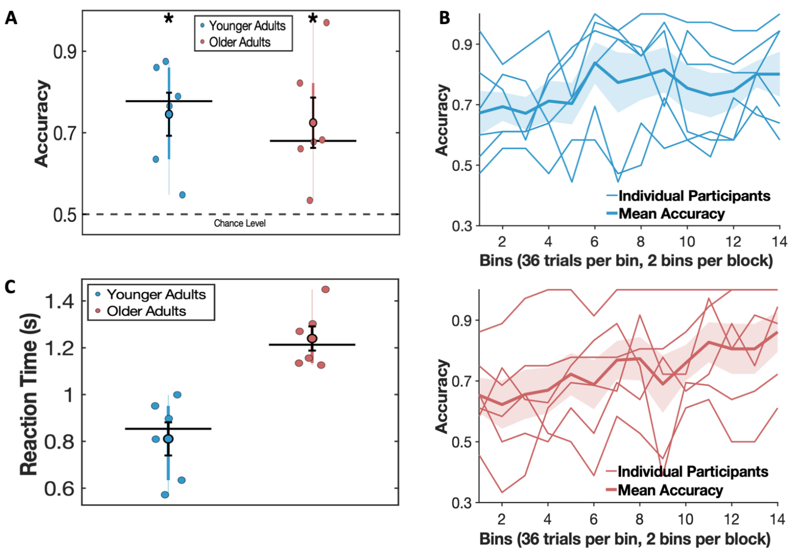

Figure 2: *Accuracy and RTs in Younger and Older Adults.* ***A.*** *Overall accuracy comparison. Boxplots illustrate the overall accuracy and its distribution in both age groups. Outlined dots with error bars represent the mean and standard error of the mean (S.E.M), while the horizontal solid line indicates the median. The horizontal dashed line marks chance level performance (0.5).* ***B.*** *Accuracy time course for younger and older adults. Accuracy was calculated for every 36 trials (a bin) to quantify the learning effect. The shaded areas represent the group average and S.E.M, while the lighter lines reflect individual participant accuracies.* ***C.*** *Average RTs comparison. Boxplots of the average RTs for younger and older adults during the choice period of standard and uncertain trials.*

To evaluate MF contributions, we examined participants' likelihood of repeating a choice on the current trial (trial n) (Prob(Repeat)) based on whether the outcome obtained from the joint outcome image was rewarded or not on the last trial (trial n-1). A purely MF agent would be more likely to repeat a choice when the joint outcome image was rewarded than when it was not. In contrast, a purely MB agent would not show this differentiation, as the effects were counterbalanced across the two counterpart bandit images sharing the same joint outcome image (see Methods for details). Consistent with Moran et al. (2019), as shown in Figure 3A, we observed that Prob(Repeat) was significantly higher for rewarded joint outcomes (J-Rew) than for non-rewarded outcomes (J-Non) in both younger and older adults (younger: $b = 1.10$, $z = 2.87$, $p = 0.004$; older: $b = 1.07$, $z = 2.63$, $p = 0.009$), suggesting a contribution from the MF system. We also found a main effect for the unique outcome (U-Rew vs. U-Non), which is predicted by both MF and MB systems (younger: $b = 2.17$, $t(474) = 4.01$, $p < 0.001$; older: $b = 1.97$, $t(496) = 3.07$, $p = 0.002$).

To quantify MB contributions, we asked whether participants generalized outcome knowledge in trials where one of the bandit images in the current trial shared a joint outcome image with the bandit chosen in the previous trial ("generalization"). A higher choice probability of the generalization bandit (Prob(Generalize)), given that the joint outcome from the last trial was rewarded, would indicate an MBCA contribution. This reflects the successful propagation of reward to a bandit not presented in the last trial, based on the transition structure. As shown in Figure 3B, MB contributions were evident in younger adults (paired-sample t-test, younger: $t(5) = -3.23$, $p = 0.023$) and marginally so in older adults (older: $t(5) = -2.15$, $p = 0.085$). However, the group difference was not statistically significant (independent t-test: $t(10) = -1.25$, $p = 0.24$). Due to the auto-correlation of the reward probability, the observed generalization effect may also reflect MF contributions. To disentangle this, we conducted a GLMM to predict choice generalization in the current trial based on the joint outcome (rewarded or not) in the last trial (Rew), its reward probability (Prob), and their interaction (Int). A significant main effect of the joint outcome (Rew) would indicate a purely MB contribution. In contrast, a significant effect of reward probability (Prob) would suggest contributions from either MF or MB systems. Our findings (Figure 3C) showed significant effects of reward probability (beta coefficient for outcome reward probability, younger: $b = 4.05$, $z = 3.53$, $p < 0.001$; older: $b = 6.42$, $z = 2.43$, $p = 0.016$) and a non-significant effect of joint outcome, indicating the MF system might entirely drive the generalization effect.

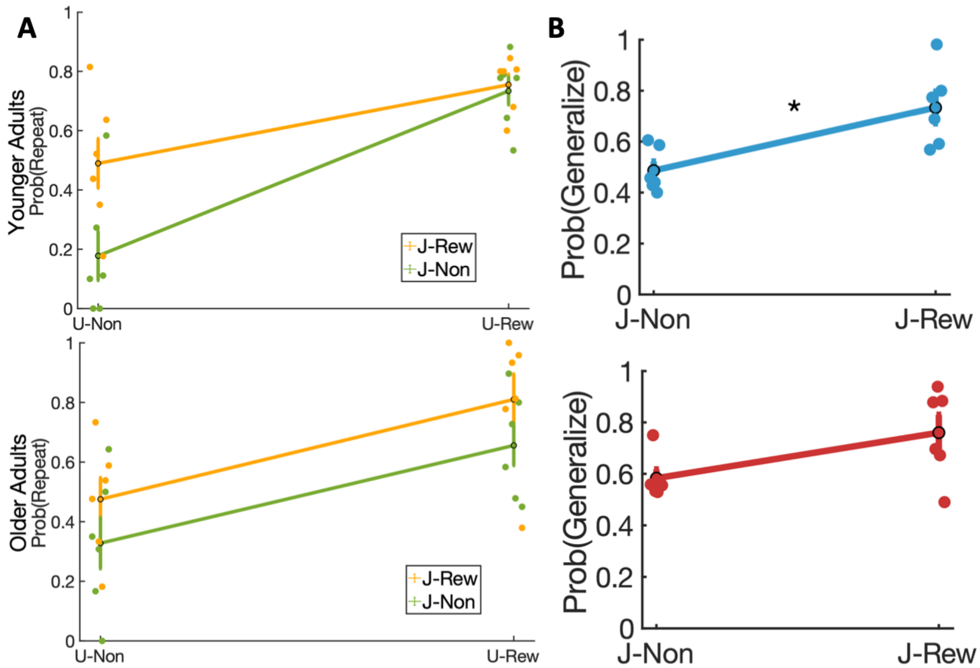

Figure 3: *MF and MB contributions to behavior.* ***A.*** *MF contribution. Standard trials featuring the same chosen bandit image as in the last trial were selected to quantify the MF contribution with a higher likelihood of repeating the same choice if it was previously rewarded (J-Rew) in comparison to non-rewarded (J-Non) in the last trial, indicating a higher contribution of MF.* ***B.*** *MB contribution. Standard trials sharing the joint outcome image with the bandit chosen in the last trial ('generalization') were selected to quantify MB contributions. This is indicated by a higher probability of selecting the generalization bandit if the shared joint outcome image was rewarded.* ***C.*** *Dissociation MF and MB contribution in the generalization choice. A positive coefficient for the joint outcome (Rew) would indicate a MB contribution. Error bars represent SEM. For A and B, dots denote individual participants.*

4 Discussion

In this pilot study, we adapted a dual-bandit task to dissociate MF and MB contributions during CA. The above-chance overall accuracy observed in both younger and older adults indicates that both groups successfully integrated the task transition structure with the dynamically changing reward probabilities to inform their decisions. Notably, older adults showed slower response times and significant learning across experimental blocks, suggesting greater cognitive effort than younger adults, consistent with previous findings (Bolenz et al., 2019; Eppinger et al., 2013; Hämmerer et al., 2019). In line with Moran et al. (2019), our preliminary results revealed significant MF contributions, evident in the increased likelihood of repeating a choice when the joint outcome image was rewarded, in both younger and older adults' choices. Although we observed a significant increase in generalization probability in younger adults and a marginal increase in older adults, it remains unclear whether the MB system was involved. This ambiguity arises from the correlation between reward probability and MF learning. As the GLMM analysis indicates, the exclusively significant effect of reward probability suggests that the MF system may have entirely driven participants' CA behavior. It is important to note that a small sample size somewhat limits our current conclusions. Overall, this pilot study provides initial evidence of age-related differences in CA and supports the feasibility of the current framework for further investigations. Future research will incorporate computational models to disentangle the contributions of the MF and MB systems and their potential dynamical interaction.

5 References

Bolenz, F., Kool, W., Reiter, A. M., & Eppinger, B. (2019). *Metacontrol of decision-making strategies in human aging. eLife, 8, e49154.* https://doi.org/10.7554/eLife.49154

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. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). *Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neuroscience, 8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Dolan, R. J., & Dayan, P. (2013). *Goals and Habits in the Brain. Neuron, 80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Eppinger, B., Walter, M., Heekeren, H. R., & Li, S.-C. (2013). *Of goals and habits: Age-related and individual differences in goal-directed decision-making. Frontiers in Neuroscience, 7*. https://doi.org/10.3389/fnins.2013.00253

Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). *Retrospective revaluation in sequential decision making: A tale of two systems. Journal of Experimental Psychology: General, 143*(1), 182–194. https://doi.org/10.1037/a0030844

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). *States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron, 66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Hämmerer, D., Schwartenbeck, P., Gallagher, M., FitzGerald, T. H. B., Düzel, E., & Dolan, R. J. (2019). *Older adults fail to form stable task representations during model-based reversal inference. Neurobiology of Aging, 74*, 90–100. https://doi.org/10.1016/j.neurobiolaging.2018.10.009

Moran, R., Keramati, M., Dayan, P., & Dolan, R. J. (2019). *Retrospective model-based inference guides model-free credit assignment. Nature Communications, 10*(1), 750. https://doi.org/10.1038/s41467-019-08662-8

Ruel, A., Bolenz, F., Li, S.-C., Fischer, A., & Eppinger, B. (2023). Neural evidence for age-related deficits in the representation of state spaces. *Cerebral Cortex (New York, N.Y.: 1991), 33*(5), 1768–1781. https://doi.org/10.1093/cercor/bhac171

Sutton, R. (1984). *Temporal credit assignment in reinforcement learning* [University of Massachusetts]. https://www.proquest.com/openview/16b17efcf37774c8c0a5e29706dc8098/1?pq-origsite=gscholar&cbl=18750&diss=y