

Exploring the neuro-computational mechanisms underlying age-related changes in  
complex decision-making

Alexa Ruel

A thesis  
In the Department  
of  
Psychology

Presented in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy (Psychology) at  
Concordia University, Montreal, Quebec, Canada

August 2023

© Alexa Ruel, 2023



## ABSTRACT

Exploring the neuro-computational mechanisms underlying age-related changes in complex decision-making

**Alexa Ruel, Ph.D**

**Concordia University, 2023**

Over the last decade, research in decision-making has made remarkable advancements in understanding how the relative engagement in model-based and model-free decision-making changes with healthy aging. While we are beginning to understand the factors that affect older adults' shift away from model-based decision-making, the exact mechanisms at play are still poorly understood. This dissertation presents findings as well as a novel theory which aims to advance our understanding of these neuro-computational mechanisms. Chapter 2 demonstrates that, in contrast to younger adults, older adults do not benefit from more distinct probabilistic transitions between stages in a two-step decision-making task. By examining trial-by-trial neurocomputational dynamics, this first empirical paper provides evidence for age-related deficits in the ability to represent probabilistic transitions, and predict the value of upcoming choice options. Chapter 3 presents a novel theory: the diminished state space theory of human aging. This theoretical contribution proposes that older adults' deficits in model-based learning are due to their underlying difficulties in representing state spaces. Chapter 4 examines one of the computational explanations brought forward in this theoretical paper. Namely, that older adults' diminished state spaces may be explained (at least in part) by their difficulties updating their internal task representation. In line with this hypothesis, results demonstrate that in contrast to younger adults, older adults show difficulties identifying outcomes that signal the need to update their internal model.

Together, these findings suggest that older adults' deficits in model-based decision-making can be explained by their diminished state space representations, which in turn may in part result from their difficulty updating their internal model during cognitive tasks. Ultimately, this dissertation provides important insights regarding older adults' deficits, and opens future directions for the study of age-related changes in representational abilities.

## ACKNOWLEDGEMENTS

First, I would like to thank my supervisor, Dr. Ben Eppinger. Thank you for the opportunity to work in your lab and for your unswaying support throughout this journey. You provided me with countless hours of help, encouragement and more scientific discussions than I can recall to which I owe much of the researcher I have become. You were and remain a great mentor and collaborator.

Second, thank you to my committee members, Dr. Virginia Penhune, Dr. Chris Steele and Dr. Karen Li who provided me with invaluable feedback and much food for thought, which I hope turn into future research questions. Of course, a huge thank you to the Natural Sciences and Engineering Research Council of Canada for supporting my research endeavors.

Thank you to all the now ex-members of the Lifespan and Decision-Making Laboratory who assisted me in more ways than I know to put in words. Sean Devine, Maxine Proffitt, Joni Shuchat, Cassandra Neumann, Nathalie Germain, Agata Kasprzyk, Kimia Motevalli, Madison Williams, Kathy Su, Jules van de Boosport, Olga Bambina and Harleen Toor. You not only made my PhD experience possible, you made it enjoyable.

Finally, thank you to my friends and family who supported and encouraged me every step of the way. You believed in me even when I didn't believe in myself. A special thank you for listening to me incessantly talk about model-based decision-making and the intricacies of collecting and analysing electroencephalography data. I am incredibly grateful to have you all in my life.

## CONTRIBUTIONS OF AUTHORS

This doctoral dissertation is composed of a general introduction (chapter 1), two empirical chapters (chapters 2 & 4), a theoretical chapter (chapter 3) and a general discussion (chapter 5).

The research question, experimental design, statistical analyses were agreed upon with my supervisor Dr. Ben Eppinger.

The sections of my dissertation were written with the feedback of my supervisor, Dr. Ben Eppinger. As mentors, Dr. Ben Eppinger and Dr. Florian Bolenz were instrumental in the theoretical and empirical sections of this dissertation. Several undergraduate students also contributed to the data collection.

Chapter 2 was published in *Cerebral Cortex* and chapter 3 is currently under review at *Perspectives on Psychological Science*. The references for the general introduction and discussion are included at the end of the dissertation, while the references for the other chapters are included at their respective ends.

The authorships are as follows:

### **Chapter 1: General Introduction**

Alexa Ruel

### **Chapter 2: Neural evidence for age-related deficits in the representation of state spaces**

Alexa Ruel, Florian Bolenz, Shu-Chen Li, Adrian Fischer and Ben Eppinger

*Ben Eppinger, Shu-Chen Li and Florian Bolenz designed the experimental paradigm and collected data. Florian Bolenz ran and interpreted the computational model. Adrian Fischer overlooked the single trial EEG regression analyses. Alexa Ruel analysed all EEG and behavioral data, interpreted these findings, and wrote the first draft. All authors reviewed and approved the final version of the manuscript.*

### **Chapter 3: Diminished State Space Theory of Human Aging**

Ben Eppinger, Alexa Ruel & Florian Bolenz

*Ben Eppinger wrote the introduction, psychological mechanisms and discussion. Florian Bolenz wrote the computational mechanisms section. Alexa Ruel wrote the neurobiological mechanisms section. All authors contributed to the creation of figures and reviewed the final version.*

## **Chapter 4: Adult Age Differences in the Updating of State Spaces**

Alexa Ruel & Ben Eppinger

*Alexa Ruel designed the novel experimental paradigm, collected, analysed and interpreted data.*

*Alexa also wrote the final manuscript. Ben Eppinger also contributed to task design, the interpretations, and reviewed the final version of the manuscript.*

## **Chapter 5: General discussion**

Alexa Ruel

## Table of Contents

<b>LIST OF FIGURES .....</b>	<b>IX</b>
<b>LIST OF TABLES .....</b>	<b>X</b>
<b>CHAPTER 1: GENERAL INTRODUCTION.....</b>	<b>1</b>
AGE-RELATED CHANGES IN DECISION-MAKING.....	2
EFFECTS OF AGING ON NEURAL REPRESENTATIONS OF TASK ENVIRONMENTS .....	5
ADVANTAGES OF USING A COMBINED COMPUTATIONAL MODELING AND EEG ANALYSIS APPROACH.....	7
CURRENT THEORIES OF COGNITIVE AGING .....	9
RESEARCH OBJECTIVES .....	11
<b>CHAPTER 2: NEURAL EVIDENCE FOR AGE-RELATED DEFICITS IN THE REPRESENTATION OF STATE SPACES .....</b>	<b>13</b>
INTRODUCTION .....	15
METHOD .....	17
DATA ANALYSIS .....	20
RESULTS .....	28
DISCUSSION .....	38
REFERENCES .....	43
<b>CHAPTER 3: DIMINISHED STATE SPACE THEORY OF HUMAN AGING .....</b>	<b>47</b>
INTRODUCTION .....	49
EMPIRICAL EVIDENCE .....	51
AGE-RELATED IMPAIRMENTS IN MODEL-BASED LEARNING RESULT FROM DEFICITS IN THE REPRESENTATION OF STATE SPACES .....	51
PSYCHOLOGICAL MECHANISMS UNDERLYING DIMINISHED STATE SPACE REPRESENTATIONS .....	53
COMPUTATIONAL MECHANISMS.....	54
NEUROBIOLOGICAL MECHANISMS.....	57
SUMMARY OF THE THEORY .....	62
SPECIFIC PREDICTIONS OF THE THEORY, GENERALIZATION AND PATHWAYS FOR APPLICATION ..	63
RELATIONSHIP TO OTHER THEORIES.....	66

REFERENCES .....	67
<b>CHAPTER 4: ADULT AGE DIFFERENCES IN THE UPDATING OF STATE SPACES</b> .....	<b>76</b>
INTRODUCTION .....	78
METHOD .....	80
DATA ANALYSIS .....	84
RESULTS .....	86
DISCUSSION .....	94
REFERENCES .....	98
<b>CHAPTER 5: GENERAL DISCUSSION .....</b>	<b>100</b>
SUMMARY AND CONTRIBUTIONS OF MAIN FINDINGS .....	100
<i>Chapter 2</i> .....	<i>100</i>
<i>Chapter 3</i> .....	<i>102</i>
<i>Chapter 4</i> .....	<i>103</i>
IMPLICATIONS .....	104
LIMITATIONS.....	107
FUTURE DIRECTIONS.....	109
CONCLUSION.....	112
REFERENCES .....	113
<b>SUPPLEMENTAL MATERIALS .....</b>	<b>122</b>
APPENDIX A: SUPPLEMENTAL MATERIAL CHAPTER 2.....	122



## List of Figures

### CHAPTER 2: NEURAL EVIDENCE FOR AGE-RELATED DEFICITS IN THE REPRESENTATION OF STATE SPACES

FIGURE 1: A) Schematic of the modified 2-stage Markov decision task. B) Trial procedure of the 2-stage task.....	19
FIGURE 2: Stay probabilities according to previous trial transition and feedback on previous trial.....	30
FIGURE 3: A) ERPs elicited by second-stage stimuli. B) Feedback-locked ERPs.....	34
FIGURE 4: Single Trial EEG Regression analysis ERPs according to corresponding PEs..	37

### CHAPTER 3: DIMINISHED STATE SPACE THEORY OF HUMAN AGING

FIGURE 1: A schematic representation of the diminished state space theory of aging.....	50
---	----

### CHAPTER 4: ADULT AGE DIFFERENCES IN THE UPDATING OF STATE SPACES

FIGURE 1: A) Schematic of the task. B) Trial procedure.....	83
FIGURE 2: Reaction time and accuracy results per trial type.....	88
FIGURE 3: ERPs elicited by feedback stimuli at Cz for each age group and corresponding topographies.....	90
FIGURE 4: ERPs elicited by feedback stimuli at Pz for each age group and corresponding topographies.....	91
FIGURE 5: Scatterplots of correlations between P3a/P3b components and accuracy.....	93

## List of Tables

### CHAPTER 2: NEURAL EVIDENCE FOR AGE-RELATED DEFICITS IN THE REPRESENTATION OF STATE SPACES

TABLE 1: Mixed-effects logistic regression on stay probabilities for first stage choices.....28

TABLE 2: Model parameters in each condition for both age groups.....31

## CHAPTER 1: GENERAL INTRODUCTION

Modern technology including cell phones and computers support many day-to-day activities including financial planning and staying in touch with family and friends. Yet older adults have consistently been shown to be slower to adapt to, and experience greater difficulties with, these technologies as compared to younger adults (Czaja & Schulz, 2006). Decades of research on the reasons for these experienced difficulties have concluded that vision loss (Gitlow, 2014), motivational changes (Tyler, De George-Walker & Simic, 2020) and computer anxiety (Di Giacomo, Guerra, Perilli & Ranieri, 2019) are all part of the problem. However, on a more fundamental level, older adults appear to have difficulties learning and making decisions within the hierarchical information structures (i.e., menus) which are inherent to most modern technologies (Ziefle & Bay, 2004; 2006). Without a clear understanding of the underlying structure of a cellphone or computer, interacting with these devices becomes difficult, and sometimes impossible (Ziefle & Bay, 2004; 2006). Consequently, many older adults avoid engaging with modern technology or rely on others to help them accomplish their digital tasks (Hunsaker et al., 2019).

Devices that were designed to allow for faster and simpler ways to stay in touch and access information may act as barriers for older adults who struggle to understand how to navigate their complex multi-level menus. This may turn out to be of particular importance for westernized societies that (1) rely more and more on these technologies and (2) have a growing proportion of older adults.

The problem is that we still know relatively little about how and to what extent decision-making abilities change during healthy aging and consequently, how these changes impact older adults' ability to engage in real-world decision-making (such as using a cellphone to send a loved one a message). My dissertation aims to address this problem by examining the neurocomputational mechanisms behind age-related changes in decision-making.

In what follows, I first provide a summary of empirical evidence for age-related impairments in model-based decision making (defined in the next section). I then review current theories of aging and highlight their shortcomings in their ability to explain recent findings in decision-making. Finally, I showcase the advantages of combining computational and

electrophysiological approaches in order to examine decision-making. The main arguments prefacing my dissertation work are that (1) we still understand relatively little about the neuro-computational mechanisms underlying age-related changes in decision-making in healthy adults, (2) that using a neuro-computational approach provides us with unique insights into the mechanisms at play, and (3) that providing a strong testable theoretical hypothesis regarding these mechanisms is necessary to better understand how decision-making changes during healthy aging.

In the following chapters, I present a first empirical study which examines the mechanisms behind age-related changes in model-based decision-making, followed by a theoretical paper which proposes a novel theory of human cognitive aging to explain these and other findings. Finally, I present a second empirical paper which aims to test one of the predictions of the theoretical manuscript.

## **Age-related changes in decision-making**

Before exploring the age-related changes in decision-making, it's important to understand reinforcement learning as its rules are intricately related to the ways in which we now examine learning and decision-making.

Inspired by the way animals and humans learn through trial-and-error, reinforcement learning (RL) is a branch of machine learning that focuses on how an agent can learn to make decisions within an environment to maximize cumulative reward (Sutton & Barto, 1999; Daw & Doya, 2006). In RL, the agent, which can be living or nonliving, interacts with an environment, receiving feedback in the form of rewards and/or punishments based on their chosen actions (e.g., lever or button presses, moves within a maze, etc.). The decision-maker typically starts off not knowing about the transitions between states (e.g., board positions in a game, locations in a maze, etc.) and the outcomes of each action. Rather, the agent develops a mapping between possible states and possible actions (policy function) over time which allow them to maximize the expected long-term cumulative reward.

Following the observation that the phasic responses of midbrain dopamine neurons in primates behaving for rewards resemble learning signals used in RL (Schultz & Dayan 1997),

neuroscientists began to integrate such RL models directly into the design and analysis of experiments on learning and decision-making (see Doll, Simon & Daw, 2012). Therefore, over the last decade, RL theories have become increasingly popular in neuroscience research as they allow us to think about and examine neural substrates for learning and decision-making in humans and animals. Because RL models are based on algorithms that describe how agents should optimally behave, they explain the mechanisms underlying observed data.

In recent years, two broad classes of RL accounts have been the focus of much attention: model-free and model-based decision-making. Learning in both model-based and model-free strategies is typically driven by prediction errors (PEs) which are defined as the difference between an actual and expected outcome. Yet each strategy relies on a different type of prediction error and therefore performs optimization in very different ways. Model-free RL for example, has been shown to rely on the reward prediction error (RPE) which tracks the difference between the actual and expected reward at a particular state (Gläscher, Daw, Dayan & O'Doherty, 2010). Therefore, model-free learning is typically defined as the process by which an agent learns simple associations between actions and rewards, often leading the decision-making agent to repeat actions that were rewarded and avoid actions that were not rewarded or that were punished (Dayan & Niv, 2008; Doll, et al., 2012). The advantage of this strategy is its efficiency and low cost in terms of cognitive resources. However, in environments where contingencies between actions and rewards change over time, model-free decision-making becomes slow due to its inflexibility. In contrast, model-based decision-making requires that the agent predict which state is currently expected given previous states and/or choices. These expectations are learned through a different type of prediction error, the state prediction error (SPE), which measures the surprise in the new state given the current estimate of action/state transition probabilities (Gläscher et al., 2010). Model-based learning therefore relies on the creation, through experience, of an internal model of the transitions and immediate outcomes in the environment. Actions that will lead to the desired outcome are chosen by searching or planning within this internal model. Engaging in a model-based strategy is much more cognitively demanding than model-free strategies but comes with the advantage of being flexible in changing environments (Doll et al., 2012).

In younger adults, choice behavior is often a mixture of model-based and model-free behavior (Daw, Gershman, Seymour, Dayan & Dolan, 2011; Gläscher et al., 2010). In contrast,

older adults have been shown to mainly rely on simpler model-free strategies (Bolenz, Kool, Reiter & Eppinger, 2019; Eppinger et al., 2013). Using a two-stage Markovian decision task in combination with computational modeling, Eppinger and colleagues (2013) found evidence for age-related deficits in model-based learning. That is, older adults demonstrated reduced model-based behavior, but no difficulties in model-free decision-making. Markov decision tasks here serve as a formal framework to represent decision-making problems under uncertainty and examine the relative contribution of model-based and model-free decision-making. Based on RL algorithms, these tasks include two decision stages, where choice at the first stage can be analysed to determine the degree to which the agent applies a model-free strategy (relying predominantly on reward information) or model-based strategy (relying on both rewards and transitions between states).

In younger adults, the decision to engage in either or strategy is often based on the benefits of doing so (Kool, Gershman & Cushman, 2017; Kool, Cushman & Gershman, 2018). In a study by Kool et al., (2017), younger adults were shown to engage in more model-based behavior when larger rewards could be obtained. Conversely, they engaged in less model-based behavior when the task became more complex as this made model-based planning more effortful (Kool et al., 2018). In contrast to younger adults, additional monetary incentives do not seem to shift decision-making behavior in older adults. Even under conditions in which engaging in a model-based strategy leads to better payoffs than model-free decision-making, Bolenz et al., (2019) demonstrate that older adults fail to show an increase in model-based behavior.

In line with previous findings that individual differences in cognitive control abilities are thought in part to reflect differences in controlled or executive-dependent processing (Kane & Engle, 2003; Otto, Raio, Chirang, Phels & Daw, 2013), Otto et al., (2013) sought to determine if model-based decision-making relies on cognitive control mechanisms such as working memory. Their results revealed that high working memory load resulted in a reduced degree of model-based behavior in younger adults suggesting that, at least in younger adults, goal-directed decisions rely on working memory functions and the associated neural systems. Based on this evidence, Eppinger et al., (2013) investigated if changes in model-based behaviour may be related to age-related and individual differences in working memory capacity. Their results replicate those of Otto et al., (2013) demonstrating that in younger adults, high working memory capacity is associated with greater model-based behavior. However, they found no association

between working memory and model-based learning in older adults, suggesting that some other factor, other than decline in working memory must explain older adults' shift towards model-free decision-making.

### **Effects of Aging on Neural representations of Task Environments**

In recent years, it has been hypothesized that older adults' shift to model-free decision-making may be due to difficulties representing the task structure of the experimental task (i.e., contingencies between actions and states in the task) (Eppinger, Heekeren & Li, 2015; Hämmerer, Müller & Li, 2014; Hämmerer, Schwartenbeck, Gallagher, FitzGerald, Düzel & Dolan, 2019); a necessary requirement for model-based learning. Specifically, early work by Eppinger et al., (2015) reveals that older adults experience difficulties learning future rewards (but not immediate rewards), indicating that older adults may have deficits in learning the transition structure of a sequential decision-making task. Similarly, using a probabilistic reversal learning paradigm, Hammerer et al., (2019) find that older adults show inconsistent choice behavior because of their overestimation of the reversal probability of the task. Consequently, they show that older adults were less able to converge on a task representation and therefore show reduced model-based behavior. Finally, Bolenz et al., (2019) found that, in contrast to younger adults, older adults showed a reduced adjustment of decision strategies as well as less reaction time slowing after changes in the task structure. In line with previous work by Decker et al., (2016), they interpret slower reaction times after surprising transitions as reflecting knowledge of the task structure, an effect that was not present in the older adults.

Taken together, these finding suggest that older adults' difficulties in representing task transition structures may be at the core of their deficits engaging in model-based learning. However, these empirical findings are unclear regarding the reason for these deficits. That is, it remains unclear if these representational difficulties arise from a difficulty in (1) creating a representation of the task, or (2) updating this representation based on trial-by-trial feedback.

From a neurobiological perspective, difficulties in the representation of the task structure are thought to result from the structural decline of hippocampal and prefrontal regions (Gershman & Daw, 2017; Lengyel & Dayan, 2007; Rajah & d'Esposito, 2005; Resnick, Lamar & Driscoll, 2007; Schuck & Niv, 2019). Beyond the well-documented role of the hippocampus

in memory (Eichenbaum, 1999), in recent years, the hippocampus has been shown to also contribute to the creation of non-spatial maps (internal task models), such as those required for goal-directed decision-making (Gershman & Daw, 2017; Lengyel & Dayan, 2007; Schuck & Niv, 2019). For instance, in a probabilistic sequential task (Bornstein & Daw, 2013), participants had to use the knowledge they gained about the likelihood of an image given the previous image. They then had to choose between two images, based on what they learned about the probabilistic relationship between images in order to maximize reward. Neuroimaging analyses revealed activation in the hippocampus, as well as in close-by cortical areas involved in the representation of the contingencies between images. In line with these findings, Kapan and colleagues (2017) demonstrated that the hippocampus is involved in maintaining representations of current and future choices during a sequential decision-making task.

Years of magnetic resonance imaging (MRI) research have observed that frontal areas such as the prefrontal and orbitofrontal cortices also experience significant deterioration with advancing age (Hedden & Gabrieli, 2004; Raz, Lindenberger, Rodrigue, et al., 2005; West, 1996). Structural changes (i.e., shrinkage) in these regions predict changes in cognitive abilities such as reversal learning (Gunning-Dixon & Raz, 2003), cognitive control (Metzler-Baddeley, Hunt, Jones, Leemans, Aggleton & O'Sullivan, 2012) and working memory (Evangelista et al., 2021). More recently, these prefrontal regions have been suggested to play a role in the representation of internal maps through helping the agent update their representation (Chan, Schuck, Lopatina, Schoenbaum & Niv, 2021; Rajah & d'Esposito, 2005; Resnick et al., 2007; Schuck et al., 2016). For instance, in the afore-mentioned paper by Eppinger et al., 2015, fMRI analyses revealed that older adults' deficits in learning the transition structure of task were reflected in their reduced dorsolateral prefrontal and orbitofrontal cortex (OFC) activity. Replicating these findings, Wittkuhn et al. (2018) showed that when repetitive transcranial magnetic stimulation (rTMS) is applied to the prefrontal cortex of younger adults, they display difficulties learning the transition structure of a task to a degree similar to older adults. More recently, using fMRI, Shuck et al., (2016) found that, during decision-making, unobservable states can be decoded from activity in the OFC, revealing that the OFC plays a role in representing a cognitive map of the task environment. Specifically, patterns of fMRI activity in the OFC contained information about the agent's current location in a mental map of the task. Therefore, the OFC seems to play a specific role in *updating* the representation of the task



environment. In line with these findings, Chan et al., (2021) demonstrate that trial-by-trial OFC activity was associated with increased expectation of the more probable outcome, indicating that the OFC plays a role in reactivation or reinforcement of a previously learned task model.

Together, these findings suggest that age-related changes in decision-making may (1) arise due to difficulties in representing the task environment, and (2) that the hippocampus and/or OFC play critical roles in these representations. What remains unclear is if older adults mainly experience difficulties in creating the internal model of the task, or in updating it based on feedback obtained during decision-making.

### **Advantages of Using a Combined Computational Modeling and EEG Analysis Approach**

Early computational research drove rapid progress in quantifying neural signals that support learning. However, much of this work was primarily focused on model-free RL (O'Doherty, 2004; Morris, Nevet, Arkadir, Vaadia & Bergman, 2006; Schönberg, Daw, Joel & O'Doherty, 2007). Although it had long been argued that the brain also uses a more sophisticated and distinct mechanism (now referred to as model-based learning) it had been much less investigated.

In a seminal paper, Daw et al., (2011) designed a novel multistep decision task to examine the relative contribution of model-based and model-free learning to behavior. Through examining choice behavior and the corresponding BOLD activity (blood-oxygen-level-dependent), they found evidence for a BOLD signal that reflects both model-free and model-based contributions. Therefore, in contrast to what they predicted, the authors found neural evidence that model-based and model-free contributions to behavior may be more integrated than previously thought. Similarly, a few years later, Deserno et al., (2015) replicated these findings showing that the neural signatures for model-based and model-free decision-making seem to overlap in the ventral striatum and PFC when examined using fMRI. It was only when researchers shifted their attention to examining the neural signatures of RPEs and SPEs, that we were able to determine that these prediction errors are computed in partially distinct brain regions, supporting the idea of overlapping, yet distinct mechanisms supporting model-based and model-free decision-making (Gläscher et al., 2010).

Altogether, while some work has provided preliminary evidence that behavioral and neural signatures for model-based learning can be identified using MRI (see also Simon & Daw, 2011;

van der Meer & Redish, 2010), most of the evidence led to the overall sense that model-based signals may not be as separable from model-free as previously thought. These findings therefore lead to the conclusion that the BOLD signal may not be precise enough to get at distinctions between model-based and model-free learning. Instead, invasive or more temporally specific methods such as EEG may be better suited.

Based on prior work suggesting that a few event-related components may reflect decision-making processes, in more recent years several researchers have used EEG to examine model-based and model-free learning mechanisms. Specifically, the P300 has been suggested to reflect the updating of the internal representation (Donchin & Coles, 1988; Nieuwenhuis, Aston-Jones, & Cohen, 2005), while the FRN (feedback related positivity) and RewP (reward positivity) have been shown to reflect feedback processing (Holroyd & Coles, 2002; Nieuwenhuis et al., 2002; Walsh & Anderson, 2012). For instance, Eppinger et al., (2017) used a combined EEG-computational RL modeling approach to examine the neural dynamics of model-based and model-free decision-making. Their results revealed that the P300 was enhanced when participants had more differentiated and reliable predictions regarding the state transitions, suggesting that the P300 may reflect SPEs. Further, the P300 amplitude correlated with the degree to which participants were able to discriminate between choice options, and the probability of choosing the best option at that point during the task. Together, in line with previous work on the P300, Eppinger and colleagues found that this component reflects the integration of model-based information about the transition structure of the task. Similarly, Nassar et al., (2019) showed that the same P300 component is associated with adjustments in learning behavior, which corroborates previous findings on its relationship to learning in changing environments (Jepma et al., 2016; Jepma et al., 2018).

Novel approaches such as multiple single-trial EEG regression analysis (see Fischer & Ullsperger 2013) provide even more insight into these mechanisms as they allow for a direct examination of the relationship between neural and behavioral data. This analysis involves using the results of the reinforcement learning model to build a general linear model (GLM) and regress single-trial EEG activity at each electrode and time point against model predictors and/or behavioral parameters. Specifically, trial-by-trial EEG data is first extracted for the time window of interest across all electrodes and all subjects and then entered into a regression with RL

computational model parameters (e.g., PEs or variable learning rates) as predictors. This first regression therefore allows for the examination of the effect of the computational model parameter on EEG signal at each electrode. In a second step, the effects of condition and age group can be added. The result is therefore a way to examine if and to what degree the relationship between computational parameters and EEG signal differs across conditions and age groups as well as possible interactions between the two. This approach, applied in study 1 (see chapter 2) provided me with novel insights into the neural mechanisms underlying age-related shifts in decision-making during aging.

### **Current Theories of cognitive Aging**

Researchers have attempted to explain cognitive aging in terms of changes in cognitive primitives, such as working memory ( Craik & Byrd, 1982), inhibition (Hasher & Zacks, 1988; Lustig, Hasher & Zacks, 2007) and speed of processing (Salthouse, 1996). While these theories have been successful in explaining general age-related changes in cognition, they fail to account for the finding that older adults show representational difficulty during decision-making.

For instance, Kirasic et al., (1996), in line with previous work by Salthouse (1991; 1996), argue that processing speed and working memory may be considered general abilities that mediate age differences in cognitive outcomes. They show that changes in working memory are a pivotal determinant of more general age-related declines in cognitive performance (Salthouse, 1995; Park et al., 2002). However, changes in decision-making appear to go beyond the age-related deficits in working memory. For instance, past work failed to find a relationship between working memory abilities and model-based decision-making in older adults (Eppinger et al., 2013). Perhaps more compelling is evidence in the same study showing that even after controlling for differences in working memory capacity, age differences in model-based behavior remained statistically significant, suggesting that working memory capacity does not fully explain age differences in model-based behavior. In line with this suggestion, in a modified predictive inference task in which participants had to infer the position of a virtual helicopter based on the positions of bags that had previously fallen from it, Nassar et al., (2016) found age differences in the ability to learn from surprising outcomes. However, these differences in learning were shown to be reflective of age rather than differences in fluid intelligence or

working memory. The authors demonstrated that when fluid intelligence and working memory were accounted as covariates in a model explaining learning differences, age was still a significant factor. Given the compelling evidence for the relationship between model-based decision-making and working memory in younger adults (see Eppinger, Walter, Heekeren & Li, 2013; Otto, Skatova, Madlon-Kay & Daw, 2015), I am not suggesting that working memory is unrelated to older adults' deficits. Rather, I propose that working memory does not appear to be the main reason older adults experience age-related deficits in model-based decision-making.

Another leading theory suggests that difficulties inhibiting irrelevant information are one major contributor to cognitive aging (Hasher & Zacks, 1988; Lustig et al., 2007). Although the work they review demonstrates that older adults have significant and important difficulties with inhibition, this deficit does not seem to significantly extend beyond interference-sensitive cognitive tasks such as the Wisconsin Card Sorting Test, the Stroop test and selective attention tasks with a predominant response. For instance, older adults have been shown to maintain intact retrieval inhibition during episodic recall (see Aslan, Bäuml & Pastötter, 2007), demonstrating that when measured with certain tasks, older adults show intact inhibition.

While these theories have been highly influential and remain critical to our understanding of how healthy aging affects cognition, as I have shown, they fail to fully explain the deficits seen during decision-making. Instead, I propose that these deficits arise from disturbances in multiple cognitive processes. Therefore, rather than relying on a single cognitive primitive, I suggest that we must develop a theory that takes multiple primitives into account in order to refine what we describe as the underlying deficit seen in aging adults' decision-making.

In line with this idea, Braver et al., (2001) have proposed the context processing theory (see also Braver & Barch, 2002; Braver, Satpute, Rush, Racine & Barch, 2005). According to their definition, context is any task-relevant information that is internally represented for task performance including representations that have their effect early in the processing stream. For example, in a Stroop task, the instructions which must be represented and maintained to guide response selection is also part of the context. Therefore, they propose that the disruption of underlying working memory, inhibition, attention, and executive function is what disrupts context processing during aging. Further, Braver et al., (2001) argue that this processing

mechanism is housed within the DL-PFC (dorsolateral prefrontal cortex) and is regulated by the DA (dopamine) system.

Although a portion of what is proposed in Braver's context processing theory is consistent with the age-related changes in decision-making described above, there are at least two inconsistencies. First, the context processing theory fails to account for several findings that older adults have deficits in representing the transition structures of decision-making tasks (Bolenz et al., 2019; Eppinger et al., 2013; Eppinger et al., 2015; Hämmerer et al., 2019). As this deficit has been shown both in probabilistic (Eppinger et al., 2013) and deterministic environments (Eppinger et al., 2015; Bolenz et al., 2019) as well as during reversals (Hämmerer et al., 2019), it cannot be ignored. Second, recent neurobiological evidence proposes that the hippocampus may play an important role in the creation of a mental model necessary for model-based learning (Gershman & Daw, 2017; Lengyel et al., 2007; Schuck & Niv, 2019). In light of the well-known effects of aging on the hippocampus, there may be a key neurobiological component missing from the context processing theory.

Overall, while current theories of cognitive aging have provided a good explanation for the age-related changes in cognition, they either explain deficits in a limited domain or in specific tasks or fail to account for recent findings in decision-making.

## **Research Objectives**

To summarize, the current literature suggests that older adults' deficits in model-based learning may arise from an underlying difficulty in representing the decision-making environment due to structural and functional changes in the prefrontal cortex (particularly the orbitofrontal cortex). However, the specific neurocomputational mechanisms by which aging adults experience difficulties engaging in model-based decision-making are unclear. Further, current theoretical accounts of aging do not fully account for existing findings on age-related deficits.

The aims of my dissertation were therefore to (1) examine neuro-computational mechanisms underlying changes in older adults' decision making, (2) determine and describe older adults' deficits in representing and/or updating their internal model and (3) to develop a theoretical framework that better accounts for age-related changes in decision-making.

In the first study of my dissertation, I examine age-related changes in representational deficits as a potential explanation for the shift towards model-free decision-making seen in older adults. By using a single trial EEG analysis approach, I directly examine the relationship between neural dynamics and behavior. In the next section, I propose a novel theory of cognitive aging which explains age-related changes in decision-making. Finally, in the last empirical paper, I examine if and to what extent older adults experience difficulties updating their internal model by examining their neural activity and behavior in a changing statistical context.

## CHAPTER 2: NEURAL EVIDENCE FOR AGE-RELATED DEFICITS IN THE REPRESENTATION OF STATE SPACES

Alexa Ruel<sup>1</sup> • Florian Bolenz<sup>2,3,4</sup> • Shu-Chen Li<sup>2,5</sup> • Adrian Fischer<sup>6</sup> • Ben Eppinger<sup>1,2,7</sup>

<sup>1</sup>Department of Psychology, Concordia University, Montreal, Quebec, Canada

<sup>2</sup>Chair of Lifespan Developmental Neuroscience, Faculty of Psychology, Technische Universität Dresden, Dresden, Germany

<sup>3</sup>Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

<sup>4</sup>Cluster of Excellence “Science of Intelligence”, Technische Universität, Berlin, Germany

<sup>5</sup>Cluster of Excellence “Centre for Tactile Internet with Human-in-the-Loop”, Technische Universität Dresden, Dresden, Germany

<sup>6</sup>Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

<sup>7</sup>PERFORM Center, Concordia University, Montreal, Quebec, Canada

Note: Copy edited version of this study was published in *Cerebral Cortex*, April 2022

### Reference

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*, 33(5), 1768-1781.

## Abstract

Under high cognitive demands, older adults tend to resort to simpler, habitual, or model-free decision strategies. This age-related shift in decision behavior has been attributed to deficits in the representation of the cognitive maps, or state spaces, necessary for more complex model-based decision-making. Yet, the neural mechanisms behind this shift remain unclear. In this study, we used a modified 2-stage Markov task in combination with computational modeling and single-trial EEG analyses to establish neural markers of age-related changes in goal-directed decision-making under different demands on the representation of state spaces. Our results reveal that the shift to simpler decision strategies in older adults is due to (i) impairments in the representation of the transition structure of the task and (ii) a diminished signaling of the reward value associated with decision options. In line with the diminished state space hypothesis of human aging, our findings suggest that deficits in goal-directed, model-based behavior in older adults result from impairments in the representation of state spaces of cognitive tasks.



## Introduction

The ability to make goal-directed decisions that rely on mental models of the environment has been shown to decline with advancing age over adult development. However, the neural mechanisms underlying these age-related deficits in complex decision-making remain unclear. In the current study, we administered a modified two-stage Markov task and used computational modeling in combination with single-trial EEG analyses to examine the neural mechanism behind the age-related shift towards model-free behavior in older adults.

Rather than making decisions guided by a mental model of the task environment (i.e., model-based decision-making), older adults tend to engage in a simpler model-free strategy, which involves learning associations between choice actions and rewards (Bolenz, Kool, Reiter & Eppinger, 2019; Eppinger, Heekeren & Li, 2015). The shift towards model-free decision-making in older adults has been hypothesized to result from age-related difficulties in representing a mental model of the task environment (Eppinger, Kray, Mock & Mecklinger, 2008; Eppinger & Kray, 2011; Hämmerer & Eppinger, 2012), a critical prerequisite for model-based decision making. These difficulties could arise from the deterioration of prefrontal brain regions during aging (Eppinger, Walter, Heekeren & Li, 2013; Raz et al., 2005; Resnick, Pham, Kraut, Zonderman & Davatzikos, 2003) as these regions are thought contribute to representing a mental model of the task environment (Schuck et al., 2016; Wilson, Fern & Tadepalli, 2014, Vikbladh et al, 2019). Aging-related declines in dopamine modulation of the frontal-striatal network (see Li & Rieckmann, 2014 for a review) may also contribute to less distinctive representations of state spaces. In line with this view, work by Wunderlich and colleagues showed that administration of L-DOPA in younger adults (to augment dopamine concentration) enhanced model-based over model-free control (Wunderlich, Smittenaar & Dolan, 2012).

In this study, our aim was twofold; a) to determine if reducing the demands on representing a mental model of the state transitions may lead older adults to engage in greater model-based decision-making, and b) to examine the neural dynamics underlying age-related deficits in goal-directed decision-making. To do so, we recorded electroencephalography (EEG) activity and adopted a single trial EEG regression approach (based on Fischer & Ullsperger, 2013; Fischer, Danielmeier, Villringer, Klein & Ullsperger, 2016). This novel approach allowed us to directly examine previous hypotheses regarding the relationship between prediction errors

and neural data for both age groups, and therefore provide unique insights into the neural mechanism underlying the age-related shift toward model-free decision-making in older adults.

To address our first aim, we had younger and older adults complete a modified two-stage Markov decision task with two conditions which differed with respect to the predictability of the state transitions (see Figure 1). In the low demand condition, common transitions between states occurred on 80% of the trials whereas rare transitions occurred on 20% of the trials. In this condition, the transition structure of the task should be simpler to represent, and thus the upcoming state should be easier to predict. In contrast, in the high demand condition, common transitions occurred on 60% of the trials whereas rare transitions occurred on 40% of the trials. In this condition the representation of the task transition structure should be less differentiated, and it should be harder for subject to predict the upcoming state. We hypothesized that both age groups would demonstrate a greater contribution of model-based decision-making in the low demand compared to the high demand condition. To address the second aim, we investigate the neural dynamics underlying age-related differences in model-free and model-based decision making and focus on two components of event-related potential (ERP) that have been shown to reflect model-based and model-free decision processes (Eppinger, Walter & Li, 2017; Sambrook, Hardwick, Wills, & Goslin, 2018): the stimulus-locked P300 component and the feedback-related negativity (FRN) respectively. In line with previous work (Eppinger et al., 2017; Gläscher, Daw, Dayan & O’Doherty, 2010), we assume that the P300 component covaries with the degree to which participants update their internal state and value representation on a trial-by-trial basis (Eppinger et al., 2017; Gläscher et al., 2010).

Next, to examine the neural mechanisms underlying feedback processing during decision-making, we focus on the FRN which has been suggested to be sensitive to negative prediction errors during reinforcement learning tasks (signed prediction errors; Holroyd & Coles, 2002; Nieuwenhuis et al., 2002; Walsh & Anderson, 2012) and more recently has been shown to reflect surprise (unsigned prediction errors; Cavanagh, Figueroa, Cohen & Frank, 2012; Talmi, Atkinson, El-Dereby, 2013). In contrast to the P300, we did not expect the FRN to vary across demand conditions based on previous work (see Eppinger et al., 2017) demonstrating that the FRN seems to reflect the contribution of a model-free strategy to decisions made at the first stage of the two-step task. However, consistent with previous work (Eppinger et al., 2008; Eppinger &

Kray, 2009; Hämmerer, Li, Müller & Lindenberger, 2011) predicted that the amplitude of FRN during probability reinforcement learning would be attenuated in old age.

## **Method**

### **Participants**

Twenty-eight healthy young adults and 30 healthy older adults participated in the study. Participants were recruited via the participant pool of the Lifelab in the Chair of Lifespan Developmental Neuroscience. We excluded participants from analysis for whom 20% of reaction times were under 200 ms (1 younger adult, 3 older adults) or who showed an insufficient understanding of the task during the practice trials, assessed by choice behavior (2 older adults). The final sample size therefore consisted of 26 younger adults ( $M_{age} = 23.73$  years,  $SD = 3.08$ , 11 males), and 25 older adults ( $M_{age} = 72.32$  years,  $SD = 3.36$ , 13 males). All participants gave informed written consent before participating. The ethics committee of the Technische Universität Dresden approved the study. Participants received a minimum payment of 5.00 euros and an additional amount (up to 8.00 euros) depending on rewards obtained within the task.

### **Stimuli**

Stimuli for the first stage of the task were two colored figures, representing the two different choice options. Stimuli at the second stage was 8 pairs of colored figures (i.e., Gogos) which were created using free software. To avoid carryover effects, 2 new pairs of stimuli were used in each block of the task. All stimuli were further processed in Adobe Photoshop. Stimuli for the second stage of the task were on a blue or brown background, representing the two different states participants could transition to at the second stage. Feedback stimuli for each trial either indicated a reward of 10-euro cents in green or 0-euro cents in red (i.e., no reward). All stimuli were presented on a 19-inch CRT computer screen using the Eprime 2.0 software (PST Inc., Pittsburgh, PA).

### **Task**

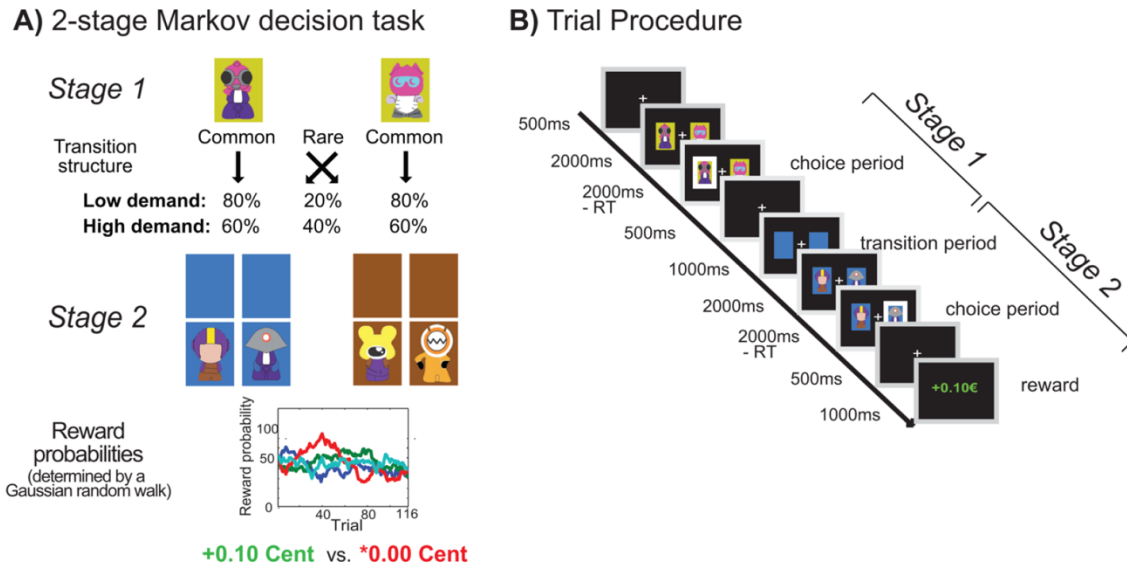
A modified version of the two-stage Markov decision-task (cf. Daw, Gershman, Seymour, Dayan & Dolan, 2011) was administered to all participants (see Eppinger et al., 2013; Eppinger et al., 2017). All participants completed two 60%–40% transition probability blocks, and two 80%–20% transition probability blocks, corresponding to a high demand and low demand condition, respectively. Each block contained 116 trials. Block order alternated within

participants and was counterbalanced across participants. Before the start of each condition block, participants were informed about the transition probability condition.

The task involved two decision stages with one decision at each stage. During the first stage, participants had to make a choice between two first-stage options (see Figure 1a). Depending on this choice, participants ended up in one of two possible second-stage states based on the transition probabilities: In the example illustrated in Figure 1, the left option is associated with a higher probability (i.e., common trials, 60% or 80%) of transitioning to the lower left second-stage state, and a lower probability (i.e., rare trials, 40% or 20%) of transitioning to the lower right second-stage state. The reverse is true for the right stimulus at stage one. Second-stage states were represented by two colored squares (with a separate color in each individual state) and upon these squares, the second-stage choice options (Gogo figures) were displayed between which participants made a choice. Finally, second-stage choices were either rewarded (10-euro cents) or unrewarded (0-euro cents) and the probability of receiving a reward followed from the chosen second-stage option. Consequently, at the second stage of the task, participants had to learn which of the 4 stimuli is currently associated with the highest probability of reward. However, to ensure participants continuously learn during the task, the reward probabilities fluctuated over time based on four independent Gaussian random walks with a standard deviation of 0.025 and reflecting boundaries of 0.25 and 0.75 (see Doll, Simon & Daw, 2012).

Choice options at both stages were randomly counterbalanced to appear on the left or right side of the screen. Responses at both stages were made by using the “f” key for the left option, and “j” key for the right option using a standard keyboard. Choices were presented for a response window of 2 s and after a choice was made, the chosen option was highlighted for the remaining time. If participants did not make a choice within the response window, the trial ended, and the task proceeded with the next trial. The colored squares signaling the second stages and the received reward at the end of the trial were each presented for 1 s. Prior to both decision stages and the reward, a fixation cross was displayed for 500 ms. To ensure all participants understood the task, a cover story was applied. The cover story was about a businessman who had to decide between two airline companies. Each airline (represented by two different figures) would bring the businessman to one of two islands on which two populations of inhabitants could be found (i.e., the Gogo figures on the blue and brown background colors). However, the airlines are somewhat unreliable regarding their destinations, and thus end up on

one island most of the time, but the other some of the time (i.e., reflecting the 80%–20% and 60%–40% transition probabilities). The task of the businessman is to make as much money as possible by tracking information about the reward probabilities of the options at the second stage of the task and the transition structure at the first stage. Importantly, participants are informed that the productivities (reward probabilities) of the populations change across time.



**Figure 1.** (A) Schematic representation of the modified two-stage Markov decision task. After making a first stage decision, participants made a second stage decision before either obtaining a reward or no reward on a given trial. At the first stage, each stimulus has a higher probability of resulting in a transition to one of the two second stage states (common transitions: 80% or 60%) and a lower probability of leading to the other of the two (rare transitions: 20% or 40%). Due to the greater predictability of the transitions in the 80%–20% condition, we predicted that both age groups should demonstrate greater model-based decision-making in this condition in contrast to the 60%–40% condition. (B) Trial procedure of the two-stage task.

## Procedure

While the experimenter prepared the electroencephalogram (EEG), participants completed a demographic questionnaire as well as the BIS/BAS personality questionnaire (Carver & White, 1994), the identical pictures, spot a word (Yuspeh & Vanderploeg, 2000) and

an N-back task. The BIS/BAS personality questionnaire was administered in order to ensure younger and older adults did not significantly differ in their relative contribution of avoidance and approach behaviors in the context of a goal-directed task. The identical pictures, spot a word and N-back tasks were administered in order to examine the relationship between verbal intelligence and working memory and performance in the two-step decisions-making task. Results of this analysis will be reported elsewhere.

Before starting the experimental task, all participants completed a computer training session. During this training session, participants were familiarized with the reward probability structure of the task and had to perform 10 choices between options with a fixed reward probability of 60%. To help their understanding of the probabilistic structure of the task, the experimenter explained the reward probabilities in terms of absolute numbers (i.e., receiving a reward in approximately six of ten trials). Participants then completed 10 trials on which they had to find the option (out of two) with the highest reward probability. After the successful completion of these 10 training trials, the experimenter explained that the reward probabilities would change slowly across the experiment. Participants were shown two examples of the random walks (Fig 1a). In the next part of the training phase, participants were introduced to the transition probabilities connecting the first and the second stage of the task. They were explained that the task involved common and rare transitions and shown an illustration of these transitions (similar to Fig 1a). Following this explanation, participants completed ten trials in which they practiced transitioning from the first to the second stage of the task. At the end of the training session, participants completed 30 full practice trials of the experimental task which involved different stimuli than in the experimental task.

## **Data Analysis**

Behavioral data were analysed in R (R Development Core Team, 2010) using the lme4 package (Bates et al., 2015). The RL model (based on Eppinger et al., 2017) and the single trial regression analyses (based on Fischer & Ullsperger, 2013; Fischer, Danielmeier, Villringer, Klein & Ullsperger, 2016) were implemented in MATLAB (Mathworks Inc. Natick, MA) using the mfit toolbox (Gershman, 2016) and EEGLAB (Delorme & Makeig, 2004), respectively.

## **First Stage Choice Data**

In an initial attempt to descriptively capture the degree to which participants engaged in model-based decision-making in the two conditions we ran a multi-level regression analysis of stay-switch behavior at the first stage of the task. Stay-switch behavior was defined as the probability of repeating a choice at the first stage as a function of the transition (common, rare) and the feedback (reward, no reward) on the previous trial. Mean stay probabilities were analysed using a mixed effects logistic regression with the factors of age group (younger or older adults), condition (60%–40% or 80%–20%), previous transition (common or rare) and previous trial feedback (reward or no reward). Only trials on which a response was generated at both the first and second stage were included in this analysis.

Based on past results (Daw, Gershman, Seymour, Dayan & Dolan, 2011; Eppinger et al., 2013; Eppinger et al., 2017), we predicted that a pure model-free decision strategy will be reflected by a main effect of reward (Figure 2). That is, that participants using a purely model-free strategy should be more likely to repeat choices that were rewarded on the previous trial and more likely to switch their first stage choice after non-rewarded trials. In contrast, a model-based decision strategy should be reflected by an interaction between the transition and the feedback received on the previous trial. This is the case because a pure model-based agent would be more likely to repeat choices after rewarded trials with a common transition and after non-rewarded trials with a rare transition. They would also be more likely to switch first-stage choices after rewarded trials with a rare transition and after non-rewarded trials with a common transition.

## **Computational Modeling**

To better understand the mechanisms underlying choice behavior in the task we ran a modified version of a well-established computational model (Daw, German, Seymour, Dayan & Dolan, 2011; Eppinger et al., 2013, Eppinger et al., 2017; Wunderlich et al., 2012). Each participant's choice behavior was fitted using a hybrid RL model. This model first acquires independent state-action values (Q values) for both a model-free and a model-based decision-making algorithm and then computes an integrated Q value as a weighted mean of model-based and model-free Q values. This weighting is controlled by a model-based weight which ranges from 0 to 1 and was held constant across trials for a given participant (but varies between

conditions). A model-based weight closer to 0 represents a greater contribution of model-free behavior, while in contrast, a model-based weight closer to 1 reflects a greater contribution of model-based decision-making. Further, the integrated Q values were converted into choice probabilities by using a softmax function (see equation 6).

The model parameters were estimated for each condition separately and participant individually as maximum a posteriori estimates. These parameters were held constant across trials but were allowed to vary across participants and within each condition across states for the same participant. Trial by trial reward prediction errors (RPEs) were extracted for each participant and were consequently used as predictors in the single trial EEG regression analyses.

### Description of the computational model

In each trial  $t$  of the task, a participant visited two states  $s_{1,t}$  and  $s_{2,t}$  at the first and second stage, respectively, decided for two actions  $a_{1,t}$  and  $a_{2,t}$  and received a final reward  $r_t$  at state 2. We model the reward expectation held for each state-action pair as  $Q(s,a)$ . Specifically, the reward expectations held by the model-free decision-making system are represented by  $Q_{MF}(s,a)$  and the reward expectations held by the model-based decision-making system are represented by  $Q_{MB}(s,a)$ .

**Model-free state-action values.** Model-free state-action values were updated using SARSA ( $\lambda$ ) temporal-difference learning (Rummery & Niranjan, 1994). After each trial, a reward prediction error  $\delta$  was computed as the difference between the expected and the actual experienced reward for each of the two decision stages.

$$\begin{aligned}\delta_{1,t} &= Q_{MF}(s_{2,t}, a_{2,t}) - Q_{MF}(s_{1,t}, a_{1,t}) \\ \delta_{2,t} &= r_t - Q_{MF}(s_{2,t}, a_{2,t})\end{aligned}$$

These RPEs were then used to update the state-action values according to the equation

$$Q_{MF}(s_{2,t}, a_{2,t}) = Q_{MF}(s_{2,t}, a_{2,t}) + \alpha_2 \delta_2, \quad (1)$$

$$Q_{MF}(s_{1,t}, a_{1,t}) = Q_{MF}(s_{1,t}, a_{1,t}) + \alpha_1 \delta_1 + \alpha_1 \lambda \delta_2 \quad (2)$$

where  $\alpha_i$  represented the learning rate at a given stage, and  $\lambda$  represented the eligibility trace decay.



Note that eligibility traces are not assumed to carry over from trial to trial because the task structure involved constantly changing reward probabilities (determined by the random walks) for each option across trials.

**Model-based state-action values.** Model-based state-action values are computed for each trial using Bellman’s equation (Sutton & Barto, 1998) by taking the model-free state-action values from the second stage and the transition probabilities into account (Eq. 3):

$$Q_{MB}(s_{1,t}, a_{1,t}) = p_1 * \left[ \max_a Q_{MF}(s_{2,t} = 1, a) \right] + p_2 * \left[ \max_a Q_{MF}(s_{2,t} = 2, a) \right], \quad (3)$$

The  $p_1$  and  $p_2$  in Eq. 3 refer to the probabilities of transitioning to each of the two second-stage states after choosing action  $a_{1,t}$  ( $p_{\text{common}}$  for common transitions  $1-p_{\text{common}}$  for rare transitions). Since transition probabilities were not explicitly instructed, we estimated  $p_{\text{common}}$  as a free parameter in both transition probability conditions.

Finally, the action-value for the full hybrid model ( $Q_{\text{net}}$ ) was calculated as the weighted sum of the model-based and model-free action-values:

$$Q_{Net} = \Omega * Q_{MB}(s_{1,t}, a_{1,t}) + (1 - \Omega) * Q_{MF}(s_{1,t}, a_{1,t}), \quad (4)$$

where  $\Omega$  is the weighting parameter. At the second stage, the  $Q_{\text{net}}$  state-action value is equal to the model-free state-action value ( $Q_{\text{net}|S2} = Q_{MF|S2}$ ).

**Softmax rule.** Choice probabilities at the first stage were calculated according to a softmax rule:

$$P_{S1}(a_1, t) = \frac{\exp(\beta_1 * [Q_{\text{net}|S1}(a_1, t) + \pi * \text{rep}(a_1)])}{(\exp(\beta_1 * [Q_{\text{net}|S1}(a_1, t) + \pi * \text{rep}(a_1)]) + \exp(\beta_1 * [Q_{\text{net}|S1}(a_2, t) + \pi * \text{rep}(a_2)]))} \quad (5)$$

where  $\beta_1$  represents the inverse softmax temperature parameter which controls the distinctiveness of the choices within each stage. We allowed both learning parameters ( $\alpha_1, \alpha_2$ ) as well as the softmax temperature parameters ( $\beta_1, \beta_2$ ) to change between both stages of the task. The indicator function  $\text{rep}(a)$  is 1 if  $a$  is a top-stage action and is the same as was chosen on the previous trial and zero otherwise. Together, the  $\text{rep}(a)$  function and the parameter  $\pi$  capture the degree of perseveration ( $\pi > 0$ ) or switching ( $\pi < 0$ ) for the first stage (Lau & Glimcher, 2005).

Choice probabilities at the second stage were calculated similarly as:

$$P_{S2}(a_1, t) = \frac{\exp(\beta_2 * Q_{\text{net}|S2}(a_1, t))}{\exp(\beta_2 * Q_{\text{net}|S2}(a_1, t)) + \exp(\beta_2 * Q_{\text{net}|S2}(a_2, t))} \quad (6)$$

The model therefore contained eight parameters ( $\alpha_1, \alpha_2, \beta_1, \beta_2, \pi, \lambda, \Omega, p_{\text{common}}$ ) and this set of parameters was estimated separately for each individual participant for each of the two conditions.

## **EEG recordings and Analysis**

Two approaches were used to examine the EEG data based on our predictions: (1) ERP analyses that allow us to study the P300 and FRN components across conditions and (2) Multiple single-trial robust regression analyses that allow us to examine the trial-wise correlation between ERP components and time-varying parameters from the computational model. Specifically, we explored the prediction that the P300 covaries with value prediction errors (VPEs) at the second stage of the task while variance in the FRN may be explained by participants' reward prediction errors (RPEs).

**Pre-processing.** EEG and electrooculography (EOG) were recorded continuously from 64 active Ag/AgCl electrodes embedded in an elastic plastic cap, using a BrainVision Recorder (Brain Products GmbH, Gilching, Germany). The electrodes were placed according to the international 10-10 system. During the recording, electrodes were referenced to the right mastoid, and re-referenced offline to the average of the left and right mastoids. The EEG signal was first filtered using a band pass filter in the range of 0.01 and 100Hz and were digitized with a sampling rate of 1000Hz. The ground electrode was placed above the forehead. Vertical and horizontal EOGs were recorded next to each eye and below the left eye. Electrode impedances were kept below 5k $\Omega$ .

For all statistical analyses, in addition to being re-sampled at 500Hz, a low-pass filter of 30Hz and a high-pass filter of 0.5Hz were applied to the EEG data. Next, bad channels were linearly interpolated and artifacts were rejected by visual inspection of the continuous data. Visual rejection resulted in 6.38% of trials being rejected for younger adults, and 5.83% of trials for older adult groups. Each data set was then epoched (-1 to 3s) surrounding the second stage choice and run through independent component analysis (ICA) to allow for further artifact rejection. Blinks, eye movements and muscle components determined using the ICA were marked and were rejected from further data analysis. All trials for which participants did not provide a first or a second stage response were removed from both the EEG and behavioral data.

**ERP Analysis.** The larger 4s epochs were then re-epoched (-200ms to 800ms) around the first stage choices, second stage stimuli and feedback phases of the task, where 0 corresponds to stimuli/feedback onset. These epochs were then baseline corrected by subtracting the average of the first 200ms pre-stimulus activity from the entire epoch (-200 to 800ms). To match trial numbers across transition probability conditions (i.e., 80%, 60%, 40% and 20%) we randomly drew a subset of trials according to the number of trials in the least frequency condition (20%) in each condition to calculate the individual participant ERP averages, resulting in 32 trials for the second stage stimulus-locked ERPs and 29 trials for the feedback-related ERPs.

***Stimulus-locked ERPs at the first stage: Choice Period.*** We analysed two ERPs at the first stage choice period: the N200 and P300 components at Pz. The N200 was measured as the mean amplitude in a 220–320ms time window after the onset of the first stage stimuli for younger adults, and as the mean amplitude in the 230–330ms time window for older adults. The P300 was measured as the mean amplitude in a 300–400ms time window after the onset of the first stage stimuli for younger adults, and as the mean amplitude in the 350–450ms time window for older adults. These times were determined separately for both age groups by building a 100ms window around the peaks which were determined by visual inspection. We then used a repeated measures ANOVA with the between subject's factor age group (younger and older adults) and the within subject's factors condition (80%–20% and 60%–40%) and transition type (common, rare).

***Stimulus-locked ERPs at the second stage: Transition Phase.*** The P300 component at the transition phase of the task was measured as the mean amplitude in a 300–400ms time window after the onset of the color patches during the transition phase for younger adults, and as the mean amplitude in the 410–510ms time window after stimulus onset for older adults. We also examined the transition phase effects at the frontal electrode FCz as the mean amplitude in a 300–400ms time window for younger adults, and as the mean amplitude in the 410–510ms time window after stimulus onset for older adults. These times were determined separately for both age groups by building a 100ms window around the peaks which were determined by visual inspection. To examine these components, we used a repeated measures ANOVA with the

between subjects factor of age group (younger and older adults) and the within subjects factors of condition (80%–20% and 60%–40%) and transition type (common, rare).

***Stimulus-locked ERPs at the second stage: Second stage Stimuli.*** The P300 component at the second stage of the task was measured as the mean amplitude in a 330- to 430-ms time window after stimulus onset (appearance of the two second stage choice stimuli on top of the colored background) for younger adults, and as the mean amplitude in the 430- to 530-ms time window after stimulus onset for older adults. These times were determined separately for both age groups by building a 100ms window around the peaks which were determined by visual inspection. To examine the mean N200 and P300 amplitudes, we used a repeated measures ANOVA with the between-subjects factor of age group (younger and older adults) and the within-subjects factors of condition (80%–20% and 60%–40%) and transition type (common, rare).

***Feedback-locked ERPs.*** The FRN was measured as the mean amplitude in the 210- to 310-ms time window after stimulus onset for younger adults, and in the 250- to 350-ms time window for older adults. These times were determined separately for both age groups by building a 100ms window around the peaks which were determined by visual inspection. For analyzing differences in the FRN amplitude, we used a repeated measures ANOVA with the between subjects factor of age group (younger and older adults) and the within subjects factors of condition (80%–20% and 60%–40%) and feedback (reward, no reward).

**Multiple Single-Trial Robust Regressions.** To examine the relationship between variables extracted from the computational model (VPE, RPE) and EEG signals we used single-trial robust regressions in a multi-level approach. We first used general linear models (GLM) to regress single-trial EEG activity at each electrode and time point against the variables extracted from the computational model. The resulting regression weights from these analyses were standardized and averaged across their respective time windows. We then used these values in set of two-way ANOVAs to examine the effects of age group and condition on the association between computational parameters and single trial EEG signals.

***Multiple Single-Trial Robust Regression of stimulus-locked ERPs at the second stage.*** To examine if VPEs explained significant variance in the P300 component elicited by second stage stimuli, we regressed trial-by-trial VPEs on the second stage stimulus-locked EEG data in each condition (60%–40% and 80%–20%) (based on Fischer & Ullsperger, 2013; Fisher et al., 2016). Regressions were run separately by condition to allow for the effect of condition to be examined at the second stage (described below). Both regressions (one per condition) were run at the subject level, resulting in a set of  $b$  values (and their associated  $p$ -values) per subject per condition. These  $b$  values therefore represent the average effect of the VPEs at each electrode in 10ms windows from -200 to 600ms. These first stage regressions revealed maximum amplitudes of positive going EEG activity at electrode Pz approximately 350ms after stimulus onset in the 60%–40% condition, and around 410ms after stimulus-onset in the 80%–20% condition which are consistent with the P300 component.

From the first level regressions, we extracted  $b$  values for each subject in each condition for the time window of interest based on the maximal effects at the first level. In order to capture the maximum amplitudes found at the first stage, we used a time window of 330–430ms. These  $b$  values from the first level regressions were standardized by their SDs before being averaged within each subject to ensure comparability between subjects. These new  $\beta$  weights representing the effect of VPE on the EEG signal were then used as the dependent variable in a set of two-way ANOVAs using condition and age group as predictors. This second level of analysis allowed for the examination of main effects of condition and age on the VPE  $\beta$  weights as well as the interaction effect between condition and age group on the VPE  $\beta$  weights. The main effects are therefore interpreted as an interaction between VPE and the main effect predictor on the P300 component, while a significant interaction is interpreted as a 3-way interaction.

In a separate analysis, we regressed the current trial's transition (common vs. rare) on the second stage stimulus-locked EEG data to examine the effect of transition on the signal in each condition. Analyses examining the effect of VPE and transition on the P300 component were run separately due to the collinearity between these predictors. All steps of analysis were identical to those described for the VPE regression and ANOVA. See *Supplemental Material* for details.

***Multiple Single-Trial Robust Regression of Feedback-locked ERPs.*** To examine if RPEs explained significant variance in the FRN component elicited by feedback, we regressed trial-by-

trial signed and unsigned RPEs on the feedback-locked EEG data in each condition. All steps of analysis were identical to those described for the VPE regressions. Using signed RPEs as predictor, first stage regressions revealed maximum amplitudes of positive going EEG activity at electrode Pz approximately 270ms after stimulus onset in the 60%–40% condition, and around 270ms after feedback-onset in the 80%–20% condition, consistent with the FRN component. We therefore used a 220–320ms time window for the second level analyses. In contrast, using unsigned RPEs as predictor, the first stage regressions revealed a significant age group effect, but no significant effects of condition. We therefore did not run second level analyses using unsigned RPEs.

## Results

### Choice Behavior

In line with previous studies (Daw et al., 2011; Eppinger et al., 2013; Eppinger et al., 2017), we examined choice behavior (stay/switch) at the first stage was analysed using condition (80%–20%, 60%–40%), previous trial transition (common, rare), the previous trial feedback (reward, no reward), as well as age group (younger adults, older adults) as predictors in a mixed-effects logistic regression (see table 1). This analysis revealed significant main effects of age group ( $\beta = -0.400$ ,  $SE = 0.107$ ,  $p < 0.001$ ) and feedback ( $\beta = 0.280$ ,  $SE = 0.046$ ,  $p < 0.001$ ). As shown in Figure 1a, older adults had a greater tendency to repeat their choices at the first stage than younger adults and across age groups participants were more likely to stick to their choice after rewarded compared to unrewarded trials. As expected based on previous studies, the analysis revealed a set of significant two and three way interaction between the factors (see table 1). Most importantly we also obtained the predicted four-way interaction between the factors age group, condition, transition and feedback ( $\beta = -0.060$ ,  $SE = 0.020$ ,  $p = 0.001$ ).

Table 1. Mixed-effects logistic regression on stay probabilities for first-stage choices.

Predictor	$\beta$	$SE$	$p$
(Intercept)	1.538	0.107	< 2e-16
Age group	-0.400	0.107	< 0.000

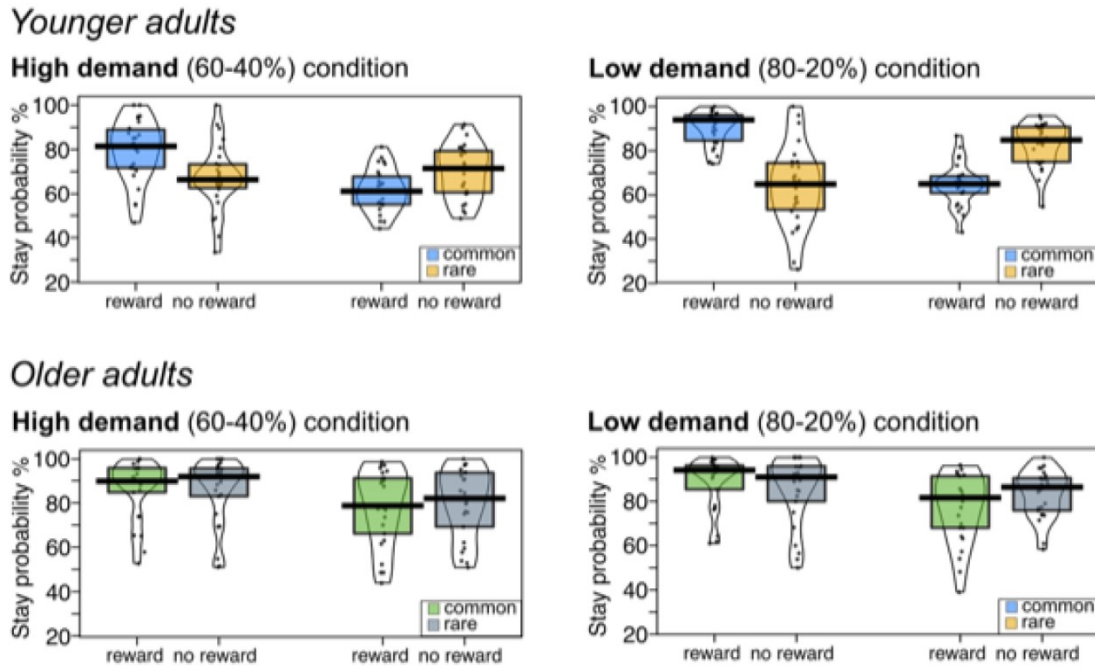
Condition	-0.110	0.043	0.010
Transition	0.054	0.029	0.060
Feedback	0.280	0.046	1.43e-09
Age group x Condition	-0.096	0.043	0.023
Age group x Transition	0.073	0.028	0.010
Condition x Transition	-0.033	0.021	0.107
Age group x Feedback	-0.061	0.046	0.186
Condition x Feedback	-0.001	0.021	0.967
Transition x Feedback	0.304	0.021	< 2e-16
Age group x Condition x Transition	-0.030	0.020	0.144
Age group x Condition x Feedback	0.020	0.021	0.330
Age group x Transition x Feedback	0.177	0.021	< 2e-16
Condition x Transition x Feedback	-0.148	0.020	3.66e-13
Age group x Condition x Transition x Feedback	-0.066	0.020	0.001

To further examine this interaction, we performed separate regression analyses for each age group. These analyses revealed significant three-way interactions between condition, transition and feedback in both age groups (*younger adults*:  $\beta = -0.215$ ,  $SE = 0.026$ ,  $p < 0.001$ , *older adults*:  $\beta = -0.081$ ,  $SE = 0.032$ ,  $p = 0.011$ ; see *Supplementary Table S1* for regression by group). To further analyse these interactions, we performed analyses separately for the two conditions and age groups.

In the 60%-40% condition, we obtained significant transition by feedback interaction for younger adults ( $\beta = 0.266$ ,  $SE = 0.031$ ,  $p < 0.001$ ) but not older adults ( $\beta = 0.047$ ,  $SE = 0.041$ ,  $p = 0.247$ ). In contrast, in the 80%-20% condition, we obtained a significant interaction between transition and feedback for both age groups (*younger adults*:  $\beta = 0.706$ ,  $SE = 0.042$ ,  $p < 0.001$ ; *older adults*:  $\beta = 0.220$ ,  $SE = 0.050$ ,  $p < 0.001$ ).

Together, these results suggest that younger adults show model-based behavior in both conditions yet relied more heavily on this strategy in the 80%–20% compared to the 60%–40% condition ( $\beta$  of 0.71 vs. 0.27). Older adults, in contrast, demonstrated no evidence of model-

based decision-making in the 60%–40% condition but did engage in the model-based decision strategy in the 80%–20% condition (see Figure 2).



**Figure 2.** Probability of repeating first choice (stay behavior) as a function of the transition on the previous trial (common or rare transition) and the feedback on the previous trial (reward or no reward). Stay probabilities are displayed separately for each condition (60%–40% and 80%–20%) across both age groups (younger and older adults). Vertical black lines represent the median, while boxes represent the inter-quartile range. Black dots represent individual participants' data, and the black outline represents the overall distribution.

### Computational Modeling Results

The analysis of estimated model parameters revealed a significant main effect of condition ( $F(1,49) = 5.288, p = 0.026$ ), a trending main affect of age ( $F(1,49) = 3.072, p = 0.086$ ) and a trending age by condition interaction ( $F(1,49) = 3.355, p = 0.073$ ) on the mixing weight parameter ( $\Omega$ ). In line with our behavioral results, this suggests that both age groups showed greater model-based behavior in the 80%-20% condition as compared to the 60%-40% condition.



In addition to the condition differences in the mixing weight parameter, we found significant age by condition interactions for the softmax ( $\beta_1$ ) ( $F(1,49) = 4.729, p = 0.035$ ) and learning rate ( $\alpha_1$ ) parameters at the first stage ( $F(1,49) = 6.738, p = 0.012$ ) as well as the  $p_{\text{common}}$  parameter ( $F(1,49) = 8.487, p = 0.005$ ). To further analyse these interactions, we performed analyses separately for the two age groups.

Analyses in younger adults revealed larger  $\beta_1$  and  $p_{\text{common}}$  parameters in the 80%-20% condition as compared to the 60%-40% condition ( $\beta_1: t = -2.044, p = 0.048$ ;  $p_{\text{common}} t = -4.589, p < 0.001$ ) whereas no such effects were obtained for older adults ( $p$ 's  $> 0.665$ ). The condition differences in the softmax parameter suggest that younger adults had a better representation of the value of the choice options than older adults did at the first stage of the task. The condition differences in the  $p_{\text{common}}$  parameter indicate that the younger adults had a more distinctive representation of the transition probabilities in the 80%-20% condition as compared to the 60%-40% condition.

Our analyses also revealed a main effect of age group for the learning rate at the second stage ( $\alpha_2$ ;  $F(1,49)=4.412, p = 0.041$ ), indicating that in younger adults recent outcomes had a greater impact on value predictions than in older adults. We also found a main effect of age group for the choice stickiness parameter, ( $\pi$ ;  $F(1,49)=17.177, p < 0.001$ ) suggesting that older adults showed more perseverative behavior than younger adults in both conditions.

Together these findings show that the pronounced differences in model-based decision-making that we observed in the descriptive analyses may not be due to underlying differences in the degree of model-basedness between age groups but may reflect differences in the representation of the transition probabilities as well as in the distinctiveness of the choice patterns at the first stage.

*Table 2.* Model parameters in each condition for both age groups.

		$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$	$\pi$	$\lambda$	$\Omega$	$p_{\text{common}}$
<b>Younger Adults</b>	<b>60%-40%</b>								
	25 <sup>th</sup> percentile	0.245	0.516	3.335	2.395	0.095	0.288	0.389	0.543
	Median	0.530	0.600	4.495	3.254	0.161	0.448	0.482	0.659

	75 <sup>th</sup> percentile	0.640	0.754	5.618	3.916	0.226	0.675	0.635	0.877
	<b>80%-20%</b>								
	25 <sup>th</sup> percentile	0.203	0.482	4.156	3.138	0.109	0.487	0.386	0.861
	Median	0.365	0.628	5.321	3.948	0.141	0.630	0.518	0.966
	75 <sup>th</sup> percentile	0.627	0.741	8.197	4.977	0.204	0.757	0.631	1.00
<b>Older Adults</b>	<b>60%-40%</b>								
	25 <sup>th</sup> percentile	0.067	0.078	4.418	1.707	0.176	0.072	0.055	0.507
	Median	0.256	0.428	4.723	2.191	0.348	0.562	0.310	0.657
	75 <sup>th</sup> percentile	0.561	0.642	6.464	3.514	0.472	1.00	0.528	1.00
	<b>80%-20%</b>								
	25 <sup>th</sup> percentile	0.337	0.082	4.332	1.719	0.181	0.245	0.332	0.512
	Median	0.470	0.532	4.848	2.122	0.256	0.951	0.539	0.658
	75 <sup>th</sup> percentile	0.730	0.942	6.145	3.265	0.471	1.00	0.654	1.00

## ERP Results

### Stimulus-locked ERPs at the first stage: Choice Period

In line with previous work (Eppinger et al., 2017), the analysis of the first stage stimulus-locked ERPs revealed a main effect of age group for the N200 component, ( $F(1,49) = 5.898, p = 0.020$ ). In addition, we found a main effect of age group ( $F(1,49) = 10.343, p = 0.003$ ), and a main effect of condition ( $F(1,49) = 5.582, p = 0.022$ ) for the P300 component. However, no significant interaction effects were for either of the components. These effects reflect that both the N200 and P300 were larger for younger adults as compared to older adults. Further, the P300 was larger in the 80%-20% condition for both age groups (see *Supplemental Figure S4*).

### **Stimulus-locked ERPs at the second stage: Transition Phase**

The analysis of the transition phase ERPs at Pz revealed only a main effect of condition ( $F(1,49) = 18.6239, p < 0.001$ ) showing that both age groups revealed a larger component (more positive going) in the 60%–40% condition ( $M = 2.730, SD = 5.924$ ) as compared to the 80%–20% condition ( $M = 1.391, SD = 4.540$ ). Transition phase ERPs at the FCz revealed a main effect of transition ( $F(1,49) = 5.422, p = 0.024$ ) as well as an interaction between transition and condition ( $F(1,49) = 5.045, p = 0.029$ ) (see *Supplemental Figure S5*).

### **Stimulus-locked ERPs at the second stage: Second stage Stimuli**

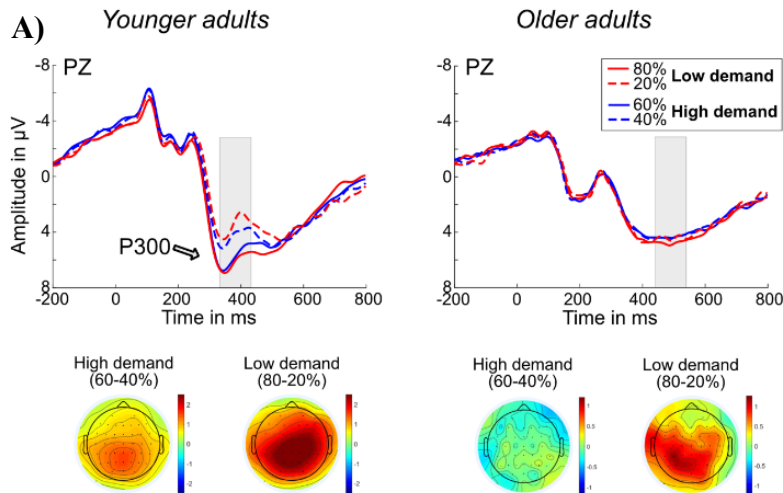
The analysis of the stimulus locked ERPs at the second stage revealed a main effect of transition, indicating that larger P300 components were elicited following common compared to rare transitions ( $F(1, 49) = 38.008, p < 0.000$ ). Further, we found significant interactions between transition and age group ( $F(2,47) = 26.377, p < 0.000$ ) as well as between transition and transition condition ( $F(2,47) = 4.296, p = 0.044$ ). To further examine the P300 effects, we ran the ANOVAs separately for the two age groups. For the younger adults, the analysis revealed a main effect of transition ( $F(1,24) = 50.073, p < 0.001$ ), as well as a significant interaction between transition and condition ( $F(1, 50) = 4.3087, p = 0.042$ ). For older adults, the same analysis revealed no significant main effects or interactions ( $p$ 's  $> 0.265$ ) (see Fig. 3A). Based on past work demonstrating that the parietal P300 component decreases with advancing age, while the frontal P300 may not (van Dinteren, Arns, Jongsma & Kessels, 2014), we ran a post-hoc ERP analysis at electrode Fz. The results revealed no significant transition by condition interaction at electrode Fz for older adults,  $F(1, 50) = 0.5564, p = 0.459$  indicating that the absence of transition effects is not due to an overall reduced P300 response at parietal sites.

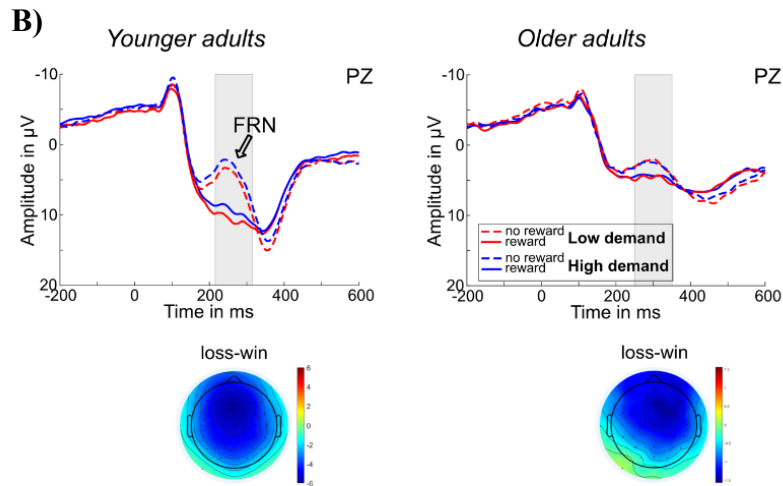
Taken together, the results show a greater P300 component in the 80%–20% condition compared to the 60%–40% condition in younger adults whereas no effect of transition probability was observed in older adults. In line with the findings from the simulated behavioral data, we interpret the larger P300 component seen in the 80%–20% condition for younger adults as evidence for an increased ability to represent the task structure in this condition. The lack of a condition effect in the older adults reveal that a more predictable environment did not help them represent the task structure.

## Feedback-locked ERPs

The analysis of the feedback-related negativity (FRN) revealed main effects of feedback ( $F(1,49) = 227.538, p < 0.000$ ), condition ( $F(1,49) = 4.670, p = 0.036$ ) and age group ( $F(1,49) = 5.561, p = 0.023$ ). Moreover, we obtained significant interactions between feedback and age group ( $F(1,49) = 57.339, p < 0.000$ ) as well as between condition and age group ( $F(1,49) = 8.144, p = 0.007$ ). To further analyse these effects, we examined each age group separately.

For younger adults, this analysis revealed a main effect of feedback ( $F(1,24) = 193.478, p < 0.000$ ) as well as a main effect of condition ( $F(1,24) = 9.084, p = 0.006$ ). These main effects demonstrate that 1) the FRN component was larger (more negative going) for no reward feedback as compared to reward feedback and 2) was more negative in the 80%–20% condition as compared to the 60%–40% condition. For older adults, this analysis revealed only a main effect of feedback ( $F(1,23) = 37.211, p < 0.001$ ) which reflects a greater (more negative) FRN on no reward compared to reward trials (see Fig 3B).





**Figure 3. A) Top:** ERPs elicited by second-stage stimuli at electrode Pz, displayed separately for the 80%–20% condition (red) and the 60%–40% condition (blue), as well as the common transitions (solid lines), and rare transitions (dashed lines) for both younger adults, and older adults. *Bottom:* The topographical map displays of the difference between common and rare transitions for the 60%–40% condition (left) and the 80%–20% condition (right) for young adults and old adults. Please note that for visualization purposes different scales are used for the topographical plots for younger and older adults. **B) Top:** Feedback locked ERPs at electrode FCz for rewards (blue) and no rewards (red) for younger adults, and older adults. *Bottom:* The topographical map displays of the differences between no reward and reward feedback across for younger adults, and older adults.

### Single Trial EEG Regression Results

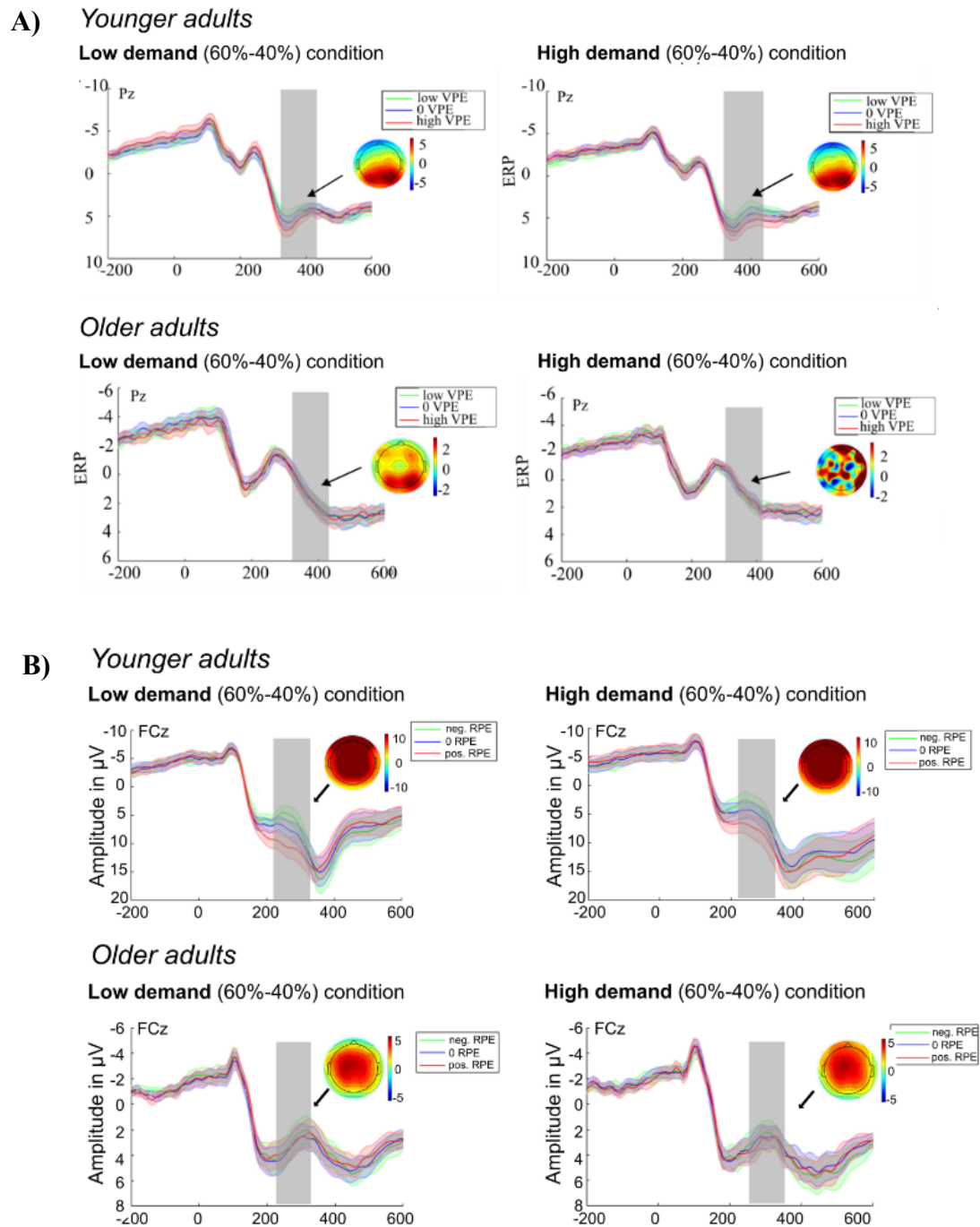
Following work by Fischer and Ullsperger (2013) and Fischer and colleagues (2016) we ran Multiple Single-Trial Robust Regression analyses on the EEG data to examine the impact of state and reward prediction errors on the corresponding event-related potentials. We then used standardized  $\beta$  weights from first the level analysis as the dependent variable in two-way ANOVAs involving the within-subject factor transition condition and the between-subjects factor age group (for details see Method section).

**VPE effects on the second stage P300 component.** The analysis of the effects of VPE's on the P300 component revealed a significant main effect of age group ( $F(1, 49) = 10.770, p = 0.001$ ) and a significant interaction between age group and transition condition ( $F(1, 49) = 5.361, p = 0.023$ ). To further examine the age group x transition condition interaction effect, we analysed each group separately.

This analysis revealed a significant main effect of transition condition for younger adults ( $F(1,49) = 4.405, p = 0.041$ ), but not older adults ( $F(1,49) = 1.188, p = 0.281$ ). As shown in Figure 4A, in younger adults VPEs explained significantly more variance in the P300 in the 80%–20% condition ( $M = 0.477, SD = 1.006$ ) compared to the 60%–40% condition ( $M = -0.037, SD = 0.739$ ). No such effect was observed in older adults.

Together, these results demonstrate that, in younger adults, the P300 component covaries with VPEs across trials and this effect is more pronounced in the 80%-20% condition in which the younger adults had a more differentiated representation of the task environment. In contrast, in older adults the P300 signal was not explained by VPEs in either condition, which points to age-related deficits in the prediction of the value of states on the second stage of the task.

**Analysis of feedback-locked EEG data.** The second level analysis of the FRN component revealed only a significant main effect of age group ( $F(1,49) = 31.685, p < 0.001$ ), indicating that the RPE explained significantly more variance in the FRN for younger adults ( $M = 0.432, SD = 1.069$ ) than older adults ( $M = -0.479, SD = 0.234$ ). As shown in Figure 4B, these findings suggest that the FRN is more tightly linked to RPE's in younger compared to older adults.



**Figure 4.** **A)** ERPs elicited by second-stage stimuli at electrode Pz of VPE variance at the second stage choice period displayed separately for young and older adults and 60%–40% and 80%–20% condition. VPEs were averaged in to high (red), neutral (blue) and low (green) values according to a tertiary split. The shaded areas within each plot represents the window of analysis (330–430ms). Topographies of maximum amplitudes are overlaid on each plot (350ms in the 6040 condition and 410ms in the 8020 condition). **B)** Feedback-locked ERPs at electrode FCz of

RPE variance at the feedback stage displayed separately for young and older adults and 60%–40% and 80%–20% condition. RPEs were averaged in to high (red), neutral (blue) and low (green) values according to a tertiary split. The shaded areas within each plot represents the window of analysis (330–430ms). Topographies of maximum amplitudes are overlaid on each plot (370ms in the 60%-40% condition and 380ms in the 80%-20% condition).

## **Discussion**

The ability to make goal-directed decisions that rely on mental models of the environment has been shown to decline with advancing age (Bolenz et al., 2019; Eppinger et al., 2013; Eppinger et al., 2015). However, the neural mechanisms underlying these age-related deficits in model-based decision-making are unclear. In the current study we used a modified version of a well-established sequential decision-making task and a novel neurocomputational approach involving model-based single-trial EEG analyses to (1) determine whether age differences in decision-making depend on the ability to internally represent the structure of the environment (the state space of the task) and (2) establish neural markers of age-related decline in goal-directed decision-making.

Younger and older adults completed a modified version of a two-stage sequential decision task (cf. Daw et al., 2011; Eppinger et al., 2013, 2017) in which we manipulated the demands on the representation of the state space in two conditions; a high representational demand condition in which the transition probabilities between the first and the second stage of the task were less differentiated (60%-40%) and a low representational demand condition in which the transition probabilities between these states were more differentiated (80%-20%) (see Figure 1).

As shown in Figure 2, the descriptive analyses revealed greater model-based behavior in both age groups in the low demand condition. In contrast, while younger adults demonstrated a significant contribution of model-based behavior in the high demand condition, in older adults we found no evidence for model-based behavior in this condition.

To better understand the mechanisms underlying these descriptive results we used a modified version of a hybrid reinforcement learning model (see Daw et al., 2011; Eppinger et al.,



2013, Eppinger et al., 2017) which was fit to task behavior. The advantage of the computational modeling approach over the descriptive analysis (which just considers choices on the first stage of the task) is that the model can be used decompose behavior across stages of the task into different assumed mechanisms. One of these parameters is the so-called Omega- ( $\Omega$ ) Parameter, which reflects degree of model-based decision-making in the task. In line with our predictions, we found a main effect of condition, indicating that both age groups demonstrated an increase in model-based decision-making in the low demand compared to the high demand condition (see Figure 2 and Table 2). Yet, in contrast to the descriptive results (and our predictions), we did not obtain a significant age group by condition interaction in the analysis in the  $\Omega$ - parameter. However, we did find expected the two-way interactions for two other parameters that are related to decision behavior at the first stage. Specifically, younger adults showed an enhanced ability to discriminate between the two first stage choices ( $\beta_1$ ) and a higher perceived common transition probability ( $p_{\text{common}}$ ) in the low demand compared to the high demand condition. No such effects were observed in older adults.

Taken together, there are two major conclusions that can be drawn from the analysis of the behavioral data. (1) Younger and older adults show greater model-based decision-making when the transition structure between states is more differentiated. This finding adds to the literature in younger adults by showing that it is not just motivational factors that support model-based decision-making (Bolenz et al., 2019) but that the complexity of the task (transition) structure matters when it comes the engagement in decision strategies that are more cognitively demanding. The fact that we also observe a shift in the degree of model-based behavior from the high to the low demand condition in older adults, may be even more noteworthy. Results from a previous study (Bolenz et al., 2019) show that in contrast to younger adults, greater incentives do not lead to more model-based behavior in older adults. This suggests that it is not older adults' willingness to engage in model-based behavior but deficits in the cognitive resources available to do so that limits them. Thus, when the cognitive demands on the representation of the task structure are reduced (such as when transition probabilities are more differentiable) older adults do show the ability to engage in greater model-based decision-making. (2) With respect to the computational mechanisms underlying age differences in model-based behavior in the two conditions, we observed an interesting mismatch between the outcome of the choice data regression analyses and the results of the computational modeling. Our descriptive regression

analysis of first stage choice behavior reveals pronounced age differences in the condition effects on model-based behavior (see Figure 2). However, these are not mirrored by a significant age group x condition interaction in the omega parameter in the computational modeling analysis. Instead, we found that younger had a better subjective representation of the transition probabilities (reflected in the  $p_{\text{common}}$ - parameter) as well as a more differentiated representations of the action values associated with the choice options than older adults. This suggests that the greater sensitivity of younger adults to the condition manipulation may not reflect greater model-based behavior per se but rather results from a greater ability to represent the transition probabilities between the first and second stage of the task.

To investigate the neural dynamics underlying age-related differences in model-based decision-making, we used two different analyses: a standard ERP analysis and multiple single-trial robust regression analyses (see Fischer & Ullsperger, 2013; Fischer et al., 2016). To our knowledge, no prior study has examined age-related changes in decision-making using single trial EEG regression analyses. This method allows us to directly examine previous hypotheses regarding the relationship between computational parameters from the model and neural data for both age groups, and therefore provide unique insights into the neural mechanism underlying the age-related shift toward model-free behavior in older adults.

For both analyses, we focused on two ERP components that have been shown previously to reflect model-based and model-free decision processes (Eppinger et al., 2017; Sambrook et al., 2018): the second-stage stimulus-locked P300 component and the feedback-related negativity (FRN).

In line with our previous results (Eppinger et al., 2017), an analysis of the stimulus-locked ERPs at the second stage of the task showed a larger P300 component in response to common as opposed to rare transitions in younger adults. Further, among younger adults the difference between common and rare transitions was larger in the 80%–20% condition compared to the 60%–40% condition. This finding points to a more differentiated representation of the task transition structure in the 80%–20% condition among young adults, which is in line both with our computational modelling results and with previous work suggesting that the P300 covaries with the degree to which participants update their representation of the probabilistic transition structure (Eppinger et al., 2017). The single-trial analyses support this view by showing that in

the younger adults, trial-by-trial changes in transition type predict changes in the second stage stimulus-locked EEG data (P300), whereas this is not the case in older adults.

In older adults the ERP analyses showed a larger P300 component for common compared to rare transitions, but no significant effect of demand condition. This finding suggests that the older adults were generally sensitive to the probability of outcomes of their choices but were unable to represent differences in probabilistic transition structures. This interpretation is consistent with our computational modelling results, showing no difference in the  $p_{\text{common}}$  parameter between demand conditions in the elderly as well as with recent findings, suggesting that older adults may have deficits in the adjustment of internal models of the task environment (Bolenz et al., 2019; Hämmerer et al., 2019).

Using single trial regression analyses, we also examined whether older adults differed from younger adults in their ability to predict the expected value of the state they ended up in on the second stage of the task (we refer to this as the state value prediction error (VPE)). VPEs were computed using Qvalues from the computational model. Therefore, these VPEs reflect participants ability to represent the relationship between their choice at the first stage of the task and the value of the options they arrive at during the second stage. As shown in Figure 5, this analysis revealed that in younger adults the P300 component covaries with the VPE and that this effect is more pronounced in the low demand condition compared to the high demand condition. Thus, younger adults were better able to predict the value of the upcoming state in the more differentiable state transition structure (i.e., in the 80%-20% condition). In older adults, we observed no such relationship between VPEs and the P300 component. In our view, the most plausible interpretation of these results is that the age-related deficits in state value predictions are a downstream consequence of diminished representations of the task structure.

Based on previous work demonstrating a relationship between the FRN component and reward prediction error processing (Eppinger et al., 2008; Eppinger & Kray, 2011; Holroyd & Coles, 2002; Nieuwenhuis et al., 2002; Walsh & Anderson, 2012), we examined age-related differences in the trial-by-trial relationship between the FRN and reward prediction errors (RPEs) during learning. Given the current debate regarding the interpretation of the FRN, we examined both signed and unsigned RPEs in order to determine if the FRN covaries with the degree to which participants recognized feedback as better or worse than predicted (signed RPE;

see Ullsperger, Fischer, Nigbur & Endrass, 2014) or the degree of surprise following feedback (unsigned RPE; Cavanagh et al., 2012; Talmi et al., 2013).

First, we investigated age differences in the feedback-related negativity (FRN) averaged across trials and as a function of transition probability and reward outcome. Based on previous results (Eppinger et al., 2017) we predicted to see a more negative FRN component following unrewarded compared to rewarded trials, but we did not expect it to covary across transition probability conditions. Consistent with these predictions, both age groups showed a larger (more negative going) FRN in response to unrewarded compared to rewarded feedback (see also Holroyd & Coles, 2002; Nieuwenhuis et al., 2002; Walsh & Anderson, 2012). Furthermore, the effect of reward outcome did not differ across transition probability conditions, which replicates our own previous results (Eppinger et al., 2017) (see Figure 4).

To examine age difference in the neural dynamics of reward prediction error processing, we performed two analyses: The first analysis used trial-by-trial signed reward prediction errors (RPEs) as the independent variable whereas the second analysis used unsigned RPEs. Interestingly, the first analysis revealed that trial-by-trial *signed* RPEs explained significant variance in the FRN for both younger and older adults in both the 60%-40% and 80%-20% condition. No such statistically significant effects were found in the second analysis when using the unsigned RPEs. Thus, our findings are in line with previous work on the FRN (Fischer & Ullsperger, 2013; Talmi et al., 2012; Walsh & Anderson, 2012), suggesting that the component reflects the processing of signed reward prediction errors during reinforcement learning.

To summarize, we find that when the cognitive demands on representing the task environment are increased, older adults tend to resort to a simpler, model-free decision-making strategy as opposed to a more demanding and flexible model-based strategy. The ERP results show that this shift in decision strategies in older adults results from deficits in their ability to 1) represent the probabilistic transition structure of the task and 2) predict the value of upcoming choice options. Thus, in line with the diminished state space hypothesis of human aging, our findings indicate that deficits in goal-directed, model-based behavior in older adults may result from impairments in the representation of state spaces of cognitive tasks.

## References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bolenz, F., Kool, W., Reiter, A. M., & Eppinger, B. (2019). Metacontrol of decision-making strategies in human aging. *Elife*, *8*, e49154.
- 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.
- Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral cortex*, *22*(11), 2575-2586.
- Feher da Silva, C., & Hare, T. A. (2018). A note on the analysis of two-stage task results: How changes in task structure affect what model-free and model-based strategies predict about the effects of reward and transition on the stay probability. *PloS one*, *13*(4), e0195328.
- 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.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.
- 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.
- 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.
- Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*(2), 521-539.
- 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*, 253.
- Eppinger, B., Heekeren, H. R., & Li, S. C. (2015). Age-related prefrontal impairments implicate deficient prediction of future reward in older adults. *Neurobiology of aging*, *36*(8), 2380-2390.

- Eppinger, B., Walter, M., & Li, S. C. (2017). Electrophysiological correlates reflect the integration of model-based and model-free decision information. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(2), 406-421.
- Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243-1255.
- Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullsperger, M. (2016). Gender influences on brain responses to errors and post-error adjustments. *Scientific reports*, *6*(1), 1-11.
- Gershman SJ. 2016. Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology* 71:1–6. DOI: <https://doi.org/10.1016/j.jmp.2016.01.006>
- 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.
- 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.
- 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*, 579-592.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, *109*(4), 679.
- Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological science*, *28*(9), 1321-1333.
- Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When does model-based control pay off?. *PLoS computational biology*, *12*(8), e1005090.
- Lau, B., & Glimcher, P. W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *Journal of the experimental analysis of behavior*, *84*(3), 555-579.
- Li, S.-C. & Rieckmann, A. (2014). Neuromodulation and aging: implications of aging neuronal gain control. *Current Opinion in Neurobiology*, *28*, 148-158.

- Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G., Holroyd, C. B., Kok, A., & Van der Molen, M. W. (2002). A computational account of altered error processing in older age: dopamine and the error-related negativity. *Cognitive, Affective, & Behavioral Neuroscience*, 2(1), 19-36.
- 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.
- Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *Journal of Neuroscience*, 23(8), 3295-3301.
- Rummery, G. A., & Niranjana, M. (1994). *On-line Q-learning using connectionist systems* (Vol. 37, p. 20). Cambridge, UK: University of Cambridge, Department of Engineering.
- Sambrook, T. D., Hardwick, B., Wills, A. J., & Goslin, J. (2018). Model-free and model-based reward prediction errors in EEG. *NeuroImage*, 178, 162-171.
- Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6), 1402-1412.
- Shahnazian, D., Ribas-Fernandes, J. J., & Holroyd, C. B. (2019). Electrophysiological correlates of state transition prediction errors. *bioRxiv*, 544551.
- Smittenaar, P., FitzGerald, T. H., Romei, V., Wright, N. D., & Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron*, 80(4), 914-919.
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol. 135). Cambridge: MIT press.
- Talmi, D., Atkinson, R., & El-Dereby, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience*, 33(19), 8264-8269.
- Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in cognitive sciences*, 18(5), 259-267.

- van Dinteren, R., Arns, M., Jongsma, M. L., & Kessels, R. P. (2014). Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. *Frontiers in aging neuroscience*, *6*, 294.
- Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., ... & Daw, N. D. (2019). Hippocampal contributions to model-based planning and spatial memory. *Neuron*, *102*(3), 683-693.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1870-1884.
- Wilson, A., Fern, A., & Tadepalli, P. (2014). Using trajectory data to improve bayesian optimization for reinforcement learning. *The Journal of Machine Learning Research*, *15*(1), 253-282.
- Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron*, *75*(3), 418-424.
- Wurm, F., Ernst, B., & Steinhauser, M. (2020). The influence of internal models on feedback-related brain activity. *Cognitive, Affective, & Behavioral Neuroscience*, *20*(5), 1070-1089.
- Yuspeh, R. L. & Vanderploeg, R. D. (2000). Spot-the-Word: A measure for estimating premorbid intellectual functioning. *Archives of Clinical Neuropsychology*, *15*(4). 319-326



## CHAPTER 3: DIMINISHED STATE SPACE THEORY OF HUMAN AGING

Ben Eppinger<sup>1,2,3,4</sup> • Alexa Ruel<sup>2,3</sup> • Florian Bolenz<sup>5,6</sup>

<sup>1</sup> Department of Psychology, University of Greifswald, Greifswald, Germany

<sup>2</sup> Department of Psychology, Concordia University, Montreal, Canada

<sup>3</sup> PERFORM centre, Concordia University, Montreal, Canada

<sup>4</sup> Faculty of Psychology, Technische Universität Dresden, Dresden, Germany

<sup>5</sup> Center for Adaptive Rationality, Max-Planck Institute for Human Development

<sup>6</sup> Cluster of Excellence “Science of Intelligence”, Technische Universität Berlin, Berlin, Germany

Note: Copy edited version of this study is currently under review at *Perspectives in Psychological Science*.

### Reference

Eppinger, B., Ruel, A., & Bolenz, F. (2023, April 25). Diminished state space theory of human aging. <https://doi.org/10.31219/osf.io/x7wpj>

## **Abstract**

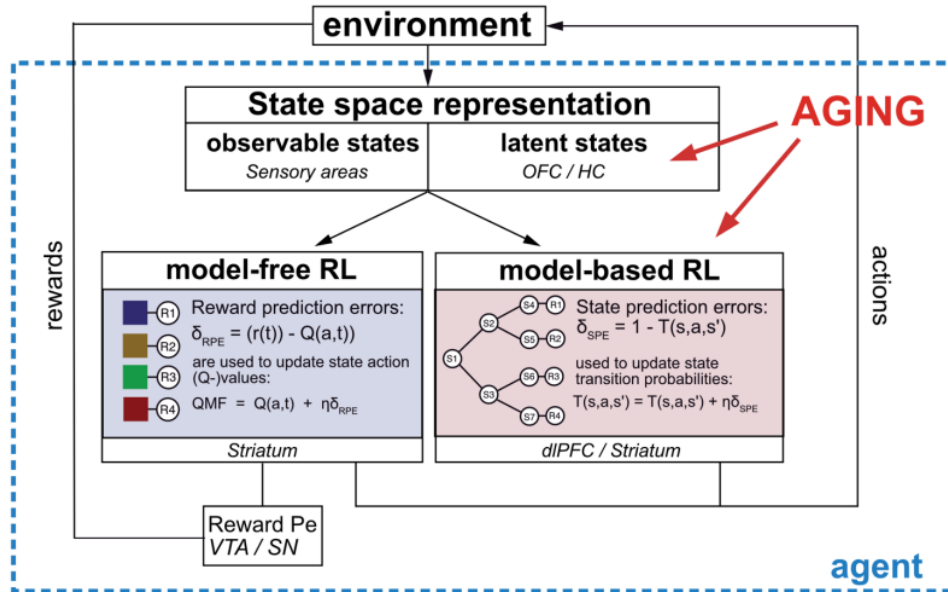
Many new technologies such as smartphones, computers, or public access systems like ticket vending machines are a challenge for older adults. One feature that these technologies have in common is that they involve underlying, partially observable, structures (henceforth called state spaces) that determine the actions that are necessary to reach a certain goal (e.g., to move from one menu to another, to change a function, or to activate a new service). In this work we provide a theoretical, neurocomputational account to explain these behavioral difficulties in older adults. Based on recent findings from age-comparative computational and cognitive neuroscience studies, we propose that age-related impairments in complex goal-directed behavior result from an underlying deficit in the representation of state spaces of cognitive tasks. Furthermore, we suggest that these age-related deficits in adaptive decision-making are due to impoverished neural representations in the orbitofrontal cortex and hippocampus.

## Introduction

The use of new technologies such as smart phones, or public access systems like ticket vending machines at train stations or check-in counters at airports, can be challenging for older adults (Czaja et al., 2006; Leung et al., 2012; Sengpiel, 2016). For example, many older adults prefer to buy train tickets at ticket counters rather than vending machines because they feel unable to navigate the complex multi-option displays of these devices (Schreder et al., 2012; Sengpiel, 2016). Thus, a technology that is meant to provide faster and more convenient access to public transport, may turn out to be a barrier for populations such as older adults. What this example means to illustrate is that many new technologies require learning and representational abilities that, with healthy aging, become more limited. This may turn out to be a problem for westernized societies that a) rely more and more on these technologies (Malik, 2014) and b) have a growing proportion of older adults (United Nations report, 2019).

In what follows, we provide a cognitive (neuro-)scientific perspective on the potential psychological and neurocomputational underpinnings of these age-related impairments in complex learning and decision-making. The main theoretical argument that we develop in this work, is that the difficulties in adaptive behavior seen in the elderly may arise from an underlying deficit in their ability to learn and represent state spaces of cognitive tasks (see Figure 1). In line with previous work (Wilson et al., 2014) we define the state space of a task as an abstract representation of the structure of the task — that is, the states that it involves, the available actions and the transitions between states that may follow from these actions (note that this definition of a state space is broader than in other parts of the literature (e.g., Sutton & Barto, 1998). In terms of the example provided in the introduction, the state space of a train ticket vending machine involves a set of menus (states) that contain choice options (actions) that link the different states (transitions to the next menu). To buy a ticket, the user (agent) must navigate and make sequential decisions in this state space. In most cases, the state space is not directly observable (latent) and must therefore be explicitly learnt from experience. In our example, a latent state could be a general setting in the vending machine that restricts the options / actions based on previous choices that have been made (e.g., discounts that have been applied). In terms of the underlying neurobiological mechanisms, we suggest that functional decline in the representational capacity of the orbitofrontal cortex (OFC) and the hippocampus (HC) are the

major source of the observed behavioral deficits (Bornstein & Daw, 2013; Bradfield et al., 2015; Kaplan et al., 2017; Schoenbaum et al., 2009; Wilson et al., 2014).



**Figure 1.** The diminished state space theory of human aging suggests that age-related impairments in goal-directed (model-based, MB) learning and decision-making result from a diminished representation of state spaces of tasks in the hippocampus (HC) and orbitofrontal cortex (OFC). In sequential decision tasks older adults primarily rely on model-free learning which has been associated with dopaminergic reward prediction error input from the ventral tegmental area (VTA) to the striatum. The dorsolateral prefrontal cortex has been shown to play a role for state prediction error signaling during model-based learning and the striatum is involved in learning by providing reward information to the MB algorithm.

## **Empirical evidence**

### **Age-related impairments in model-based learning result from deficits in the representation of state spaces**

Recent computational work on reinforcement learning in humans has focused on two different decision strategies: model-free decision-making, which relies on the learning of simple associations between states or actions and reward, and model-based decision-making, which relies on the learning of a state space (model) of the environment that is used for forward planning of actions (Daw et al., 2011; Gläscher et al., 2010). The advantage of model-free decision-making is its computational efficiency and low cost in terms of cognitive resources. The downside of this strategy is that it is inflexible and slow in adapting to changes in the environment because the agent has to relearn the contingencies from experience. In contrast, the advantage of a model-based strategy is its flexibility: If rewards or contingencies between states in the environment change, the learner can quickly adapt behavior by updating the state space representation.

Empirical studies have used variants of so-called multi-stage Markovian learning tasks to study the interplay of model-free and model-based decision processes. A prominent version of such a paradigm is the so-called two-stage decision task developed by Daw and colleagues (2011). In this task, a decision at the first stage (state 1) determines a probabilistic transition to the second stage, which contains two states. Participants then have to decide between two options and receive a probabilistic reward. The participants are supposed to learn what the option with the highest reward is at a given time and how to transition from one state to the other in order to accumulate as many rewards as possible. The choice data generated through this task can be used to estimate the degree of model-based behavior for each individual. That is, how well they integrate their representation of the state space with knowledge about the reward probabilities in order to accumulate as much reward as they can.

In younger adults, choice behavior in this task reflects a mixture of model-based and model-free decision processes, but there are substantial individual differences in the degree of model-based control (Daw et al., 2011). For example, it has been shown that individuals with higher working memory capacity engage in greater model-based decision-making and are less susceptible to the detrimental effects of stress on this strategy (Eppinger, Walter, et al., 2013b;

Otto et al., 2013; Schad et al., 2014). Furthermore, in younger adults higher cognitive control abilities are associated with greater model-based decision-making, indicating that the two behaviors may share a common set of underlying computational processes (Otto et al., 2015). Recent work also points to an association between the ability to infer latent structures in the environment and model-based control in the two-stage task (Markant, 2020; Rmus et al., 2019). Together, these results suggest that model-based control relies on the ability to represent and update an internal model of the task structure (the state space).

In contrast to younger adults, older adults rely predominantly on model-free learning strategies and show performance deficits under conditions that require model-based learning (Bolenz et al., 2019; Eppinger et al., 2015; Eppinger, Walter, et al., 2013a; Ruel et al., 2022). Furthermore, findings from these studies suggest that the degree to which older adults shift towards model-free strategies depends on their ability to represent the state space of the task. Results by Bolenz et al. (2019) show evidence for an adjustment of decision strategies after changes in the task structure (the state space) in younger adults. Furthermore, these adjustments were associated with reaction time slowing, indicating that the behavior reflects a deliberate adaptation to changes in the state space. In contrast, older adults show deficits in the updating of task representations as well as less reaction time slowing after changes in the task structure, supporting the idea of deficits in state space representations. In a recent study, Ruel et al. (2022) investigated age differences in model-based decision-making under varying demands on the representation of the task structure. Consistent with the data by Bolenz et al. (2019), the behavioral results show substantial age-related deficits in model-based decision-making under high demands on the representation of the task structure. Even under low representational demands, older adults show reduced model-based decision behavior when compared to younger adults. Similar findings come from a study on age differences in reversal learning, which showed that older adults are less able to converge on stable task representations and that this deficit is independent of age differences in outcome processing (Hämmerer et al., 2018). It is important to note that although the sequential decision tasks applied in the studies outlined above are similar in that they can be described as a Markov decision process, they also differ in many important respects: The experimental tasks used in Eppinger et al. (2013) and Ruel et al. (2022) involved probabilistic transition structures whereas in the study by Bolenz et al. (2019) deterministic structures were applied that involved reversals of state transition mappings. The studies by

Eppinger et al. (2015) and Wittkuhn et al. (2017) applied deterministic transition structures but across three instead of two states and with rewards as well as punishments. In the study by Bolenz et al. (2019), performance on the task (reflected in monetary payouts) depended on the degree to which participants engage in model-based behavior whereas in the studies by Eppinger et al. (2013) and Ruel et al. (2022) performance was independent of the degree of model-based behavior. Thus, the state space theory is partly based on an empirical generalization, but the experimental designs of the underlying tasks vary substantially, which allows for a detailed examination of the deficits experienced by older adults.

That being said, we do not want to imply that age-related differences in learning and decision-making are only due to a diminished representation of state spaces. For example, results of several studies indicate that older adults might have difficulties in building up reward value representations, particularly under high degrees of reward uncertainty (Chowdhury et al., 2013; Eppinger et al., 2011; Eppinger, Schuck, et al., 2013; Lerner et al., 2018). Yet, these deficits do not seem to be the primary sources of the shift from model-based to model-free control that is observed in healthy aging. As shown in Eppinger et al. (2013) more differentiated reward structures do support model-based behavior in younger adults, but not older adults.

### **Psychological mechanisms underlying diminished state space representations**

A straightforward explanation for the observed age-related deficit in model-based learning might be the well-documented age differences in working memory capacity. Findings by Otto et al. (2013) show a significant positive association between working memory and model-based control in younger adults. Results from Eppinger et al. (2013) replicate these findings in the young. However, they show no such correlation in older adults. Furthermore, when controlling for differences in working memory capacity, age differences in model-based behavior remained statistically significant, suggesting that working memory capacity does not fully explain age differences in model-based behavior. Again, this is not to say that model-based control is independent of working memory, as the evidence for such a positive relationship between model-based behavior and working memory in young adults is compelling (Eppinger, Walter, et al., 2013a; Otto et al., 2015). However, it seems that working memory may not be the primary reason for the limitations in model-based behavior observed in older adults.

An alternative (although not mutually exclusive) explanation for the shift from model-based to model-free behavior with age might be age-related deficits in the consolidation of state space information (Schuck & Niv, 2019). Using a sequential decision task in combination with multivariate pattern analyses (MVPA) of fMRI data, Schuck and Niv (2019) show evidence for a sequential replay of task states in the hippocampus. Furthermore, these replay mechanisms seem to drive the learning of complex state spaces in other brain areas such as the orbitofrontal cortex (see also (Momennejad et al., 2018)). Accumulating evidence suggests that replay also occurs in cortical areas and can be decoded from fMRI resting state data (Wittkuhn & Schuck, 2021). Work by Liu et al. (2021) supports this view by showing a backward cortical replay signal in MEG data that is associated with learning performance. They suggest that this replay signal may provide a mechanism for solving non-local credit assignments problems during model-based learning by connecting actions and outcomes across intervening states (Liu et al., 2021).

To summarize, it seems very likely that the replay of state sequences is critical for the construction and use of state space representations. There is indirect evidence from the aging and episodic memory literature that shows a reduced association between neural reinstatement signals in the medial temporal lobe and memory performance in older adults (Stawarczyk & Wahlheim, 2020). However, to our knowledge, the direct relationship between the replay signals described above and learning and memory abilities in older adults has yet to be examined. In the following two sections we will consider potential computational implementations of the diminished state space theory and consider how impoverished state space representations might be reflected in the brain.

## **Computational mechanisms**

In computational reinforcement-learning models, a state space can be represented by a set of states in the environment, a set of possible actions and the transition probabilities  $P(s'|s,a)$  that specify the subjective assumptions about how likely performing action  $a$  in state  $s$  will lead to the new state  $s'$ . By means of these transition probabilities, the expected reward  $Q(s,a)$  for a state action pair can be computed as the weighted mean of the reward in subsequent states. The reward of a subsequent state is composed of the reward immediately available in this state  $r(s')$



and the temporally discounted expected future reward from choosing the most valuable action in this state.

$$Q(s, a) = \sum_{s'} P(s'|s, a)[r(s') + \gamma \max_{a'} Q(s', a')]$$

Note the recursive nature of this computation: To evaluate the potential consequences of an action, not only do the immediately subsequent states have to be considered, but also all possible subsequent states after two, three, or more steps. Thus, computing a reward expectation is not only guided by the part of the state space representing the transition probabilities from the current state, but requires a representation of the transition probabilities for all states that can potentially be reached at some point in the future. We propose that the diminished state space theory of human aging could be reflected in two different ways in computational models. First, as an incomplete or inaccurate representation of states, actions, or transition probabilities and, second, as an inefficient updating of these transition probabilities.

#### *Simplified representations of the state space*

Older adults might have difficulties with maintaining information about differences in transition probabilities (e.g., “this action will lead with 80% probability to state X and with 20% probability to state Y”). Therefore, rather than trying to accurately represent the state space, they may only maintain information about the states that are connected and consider all these connections to be equally likely (e.g., “this action will lead to either state X or state Y with equal probability”). The idea that older adults might simplify their representation of transition probabilities receives support from the study by Ruel et al. (2022), which showed that older adults represented transition probabilities as almost equally likely even when in fact one transition was five times more likely than the other possible transition. Alternatively, it could be that older adults simplify their representation of the state space by pruning branches of the decision tree with aversive outcomes or with low frequencies, thus reducing the number of states or actions. Similarly, older adults might restrict the planning depth of their state space, for example, by only representing information of immediately subsequent states but not of states thereafter (see Huys et al., 2012; Huys et al., 2015; Keramati et al., 2016 for findings on these strategies in younger adult samples). Findings also show that humans can use successor representation learning, a strategy that relies on stored predictions of future states (instead of

transition probabilities between all possible states) and is therefore simpler than a model-based strategy while preserving a lot of its flexibility (Momennejad et al., 2017). Finally, recent work by Ho et al. (2022) suggests that people adaptively construct simplified representations of cognitive tasks when planning action sequences (Ho et al., 2022). The authors refer to these as task construals, and it is possible that older adults are less flexible in adjusting the complexity of these representations to task demands. Taken together, although there is considerable evidence for simplified representations of state spaces in older adults, the mechanisms that govern these simplifications are not clear yet and need to be established in future studies.

### *Inefficient updating of transition probabilities*

One of the major benefits of model-based decision strategies is that they can flexibly adapt to changes in the task structure. However, in order to do so, these changes need to be incorporated by updating the transition probabilities. Previous computational accounts have modeled this with state prediction errors and transition learning rates similar to temporal-difference learning (Bolenz et al., 2019; Gläscher et al., 2010). The magnitude of the transition learning rate then determines the degree to which experienced transitions are disregarded as oddballs or random perturbations (learning rate = 0) or interpreted as indicators for changes in the underlying task structure (learning rate close to 1). There is no universal value for an optimal transition learning rate. Rather, the transition learning rate has to be flexibly adjusted based on the statistical properties of the environment (D'Acremont & Bossaerts, 2016; Nassar et al., 2019). Age-related deficits in the flexible adjustment of transition learning rates might lead to an insufficient (rigid) representation of the state space, consistent with the predictions of the diminished state space theory. However, age differences in model-based behavior are also observed when transition probabilities are deterministic (Bolenz et al., 2019; Eppinger et al., 2015), which indicates that diminished state space representations are not due to deficits in learning probabilistic state transitions alone. One reason for an insufficient updating of learning rates may be age-related alterations in episodic memory replay that is necessary for keeping track of the changes in state transitions (Olafsdottir et al., 2018; Schuck & Niv, 2019). Previous work suggests a role of reward prediction errors in prioritizing replay of state transition sequences (Momennejad et al., 2018). However, whether state prediction errors have a similar effect in supporting replay is currently unclear. Another reason for a reduced updating of transition

probabilities in older adults could be that suboptimal transition learning rates in older adults reflect a greater tendency to perseverate on previous predictions in the elderly (Nassar et al., 2016). A greater rigidity of internal state spaces might be beneficial in some contexts (Devine et al., 2022) but is mal-adaptive in environments in which performance depends on a flexible adjustment of state space representations. To summarize, each of these computational explanations seems consistent with what is seen behaviorally (Bolenz et al., 2019; Hämmerer et al., 2018; Ruel, Bolenz, et al., 2021). Yet, the exact mechanism that leads to these deficits in updating is unclear. Overall, we have presented two potential computational explanations for diminished state space representations in older adults: (1) Simplified representations of state spaces and (2) an inefficient updating of state transition probabilities. It is important to note that these explanations are neither mutually exclusive nor completely interchangeable. That is, older adults might have a simplified representation of the task structure even when no updating of the transition probabilities is needed (e.g., stable-transitions condition in Bolenz et al. (2019)). Similarly, if older adults are simply slower at incorporating state prediction errors in their internal model (inefficient updating), they should nevertheless eventually converge on the same mental representation of the task as younger adults.

### **Neurobiological mechanisms**

Several recent studies tried to establish the neural processes underlying goal-directed decision making. Results of this work suggest that the orbitofrontal cortex (Nogueira et al., 2016; Schuck et al., 2016; Wilson et al., 2014) and the hippocampus (Gershman & N.D., 2017; Lengyel & Dayan, 2007; Schuck & Niv, 2019; Stachenfeld et al., 2017) play key roles for the representation and updating of state spaces during model-based decision-making. In what follows, we go over recent findings on the role of each of these areas in goal-directed behavior and discuss how aging might affect the neural mechanisms underlying model-based learning and decision-making.

**Hippocampus.** Seminal findings by O’Keefe and Nadel (1978) demonstrated that activity in the rat hippocampus reflects the use of cognitive maps during spatial navigation (O’Keefe & Nadel, 1978). More recent work by Johnson and Reddish (2007) showed that these hippocampal

representations are not static, but are updated as the animal moves along its path in the maze, suggesting that spatial navigation is an active process that is likely controlled by other cognitive mechanisms (Johnson & Reddish, 2007). Situated in the medial temporal lobe, the human hippocampus appears to play a similar role as the rat hippocampus by learning locations relative to a spatial boundary (Doeller et al., 2008). However, in humans the hippocampus has also been associated with more abstract (i.e., not exclusively spatial) representations of episodic information (Deuker et al., 2016; Shohamy & Wagner, 2008; Tavares et al., 2015). These findings led to the hypothesis that the hippocampus may play a role in representing state space information necessary for model-based reinforcement learning (Gershman & Daw, 2017; Lengyel & Dayan, 2007; Schuck & Niv, 2019). More specifically, the hippocampus has been suggested to be involved in the representation and planning of sequential actions necessary to achieve a certain goal. To gain a better understanding of how the brain learns these sequential state-action contingencies, Bornstein & Daw (2013) had participants view a series of images in which the next image to be presented depended probabilistically on the current image that they observed. At test, participants were asked to use the knowledge they had gained in the first phase of the task (regarding the sequential contingencies between stimuli) to guide their choice between two images in order to maximize reward (Bornstein & Daw, 2013). The neuroimaging analyses revealed that the hippocampus and close-by cortical areas are involved in the representation of state contingencies during model-based decision making. Work by Kaplan and colleagues (2017) echoes these findings by demonstrating that in a spatial sequential choice task the hippocampus is part of a network which supports planning by maintaining representations of current and potential future choices (Kaplan et al., 2017). More recently, in a neuropsychological study, Vikbladh and colleagues (2019) found that the degree to which participants' planning was impaired was related to the amount of damage to the right hippocampus (Vikbladh et al., 2019). Further, in line with findings in rodents, recent work has shown that sequential replay of experiences in the hippocampus and cortex during decision-making also occurs in humans (Liu et al., 2019; Liu et al., 2021). Overall, it seems that the hippocampus may be involved in representing two important pieces of information necessary for model-based planning and decision-making: a) the transition probabilities between different states and b) the states that can be potentially reached in the future. What remains unclear from

the current work on the involvement of the hippocampus in goal-directed planning and choice behavior is the degree to which it can be extended beyond the spatial domain.

**Orbito-frontal Cortex.** Another brain region proposed to be involved in goal-directed decision-making is the orbito-frontal cortex (OFC). The OFC is located on the ventral surface of the prefrontal cortex and has been previously suggested to play a key role in the representation of reward value, or more generally, in hedonic experience (Kringelbach, 2005; Rolls, 2000; Wallis, 2007). However, recent work points into a different direction and suggests that the OFC is involved in the representation of an agent's current location within an abstract cognitive map of a task (Bradfield et al., 2015; Schoenbaum et al., 2009; Wilson et al., 2014). Recent empirical evidence provides support for this theory. For instance, Schuck et al. (2016), using multivariate pattern analysis, found patterns of activity in the OFC during a 16-state decision-making task that contained unobservable information about the participants' location in a cognitive map of the task. Performance on the task depended on memory of past events and knowledge about the current phase of the task. In line with previous animal studies (Takahashi et al., 2011), the authors suggest that the human OFC may represent information about locations in a mental (i.e., non-spatial) map of a task that is comprised of hidden or partially observable states. More recently, Chan et al. (2021) tested this prediction empirically. Using univariate and multivariate activity fMRI analyses, they demonstrated that activity in the OFC correlates with experienced state transitions. Their findings reveal that the OFC is also involved in learning a state-to-state transition structure, necessary for accurate planning during model-based decision-making (Chan et al., 2021).

What about the interplay of the OFC and HC during decision-making? Several findings indicate that the hippocampus and OFC support separate functions such as episodic memory and the representation of subjective (reward) value. However, there is also evidence that points to complementary functions when it comes to decision-making. Work by Wang, Schoenbaum and Kahnt (2020) indicates that the OFC and hippocampus represent partially overlapping information such that the HC may be involved in the initial acquisition of information for the creation of a cognitive map whereas the OFC seems to maintain these cognitive maps and to guide behavior (Wang et al., 2020) (see also (Schuck & Niv, 2019; Wimmer & Büchel, 2019)). Further, work by Johnson et al. (2022) suggests that the OFC, similarly to the hippocampus,

supports an understanding of temporal order information (Johnson et al., 2022). In line with previous work by Barbey et al. (2011), these findings highlight the role of the OFC in working memory and temporal cognition (Barbey et al., 2011). Coming back to the train ticket example in the beginning, while the hippocampus appears to help the user represent the overall relationship between states, the OFC may play a role in keeping track of the states they are currently in (i.e., menu) based on previous actions (i.e., choices) and previously visited states. However, the degree to which the roles of these two structures overlap is poorly understood.

Taken together, current findings indicate a role for the OFC as well as hippocampus in creating, maintaining and updating non-spatial state space representations during goal-directed decision-making. Specifically, these findings suggest that the OFC and HC represent partially overlapping information and that interactions between these two regions support model-based planning and decision-making. While the hippocampus has been suggested to be involved in the creation of a “cognitive map” of the state space and the transitions between them, the OFC appears to help the agent use this map to guide behavior.

### **The effects of aging on hippocampal and OFC function**

A considerable body of structural MRI work indicates that prefrontal regions and the hippocampus deteriorate as we age (Hedden & Gabrieli, 2004; Raz et al., 2005; West, 1996). In line with the structural findings, functional MRI research suggests an under-recruitment of various areas in the prefrontal cortex including the OFC (Rajah & d’Esposito, 2005; Resnick et al., 2007). This under recruitment of the PFC, and particularly the OFC, could limit top-down activation of the model-based learning network leading older adults to rely more on simpler, model-free decision strategies. According to the neurobiological evidence, a shift away from model-based learning could be related to changes in the function of: (1) the hippocampus due to its involvement in the representation of state spaces and/or (2) the orbitofrontal cortex in line with its role in helping the agent use and update these representations.

Research on age-related changes in memory has long shown that healthy aging is associated with a decrease of contextual (episodic) memory, while memory for content (semantic information) remains more intact (see (Spencer & Raz, 1995). To our knowledge the relationship between age-related changes in hippocampal structure and/or function and model-based behavior has not yet been established. Nevertheless, based on the findings in younger adults, hippocampal

decline may lead to difficulties representing state transitions required for a complete “cognitive map” of the state space (Bornstein & Daw, 2013). Further, based on findings by Liu et al., (2019, 2021) demonstrating that hippocampal and cortical replay aids the creation of a mental model of the decision-making environment as well as model-based learning, age-related decline of the hippocampus may contribute to changes in decision-making during aging. Based on the neuropsychological findings by Vikbladh et al. (2019), we would predict that the degree of hippocampal deterioration would affect both decision-making and spatial navigation.

Beyond the effects on the hippocampus, healthy aging also strongly affects the structure and function of prefrontal regions (West, 1996). Specifically, with increasing age, various prefrontal regions appear to be underrecruited, which may lead to difficulties in older adults in using and updating their state space representation. Findings from an fMRI study by Eppinger et al. (2015) support this view showing that impairments in older adults in learning to predict future rewards are associated with prefrontal deficits in extracting sequential state transition structures (Eppinger et al., 2015; Wittkuhn et al., 2018). A study by Wittkuhn et al. (2018) replicated the behavioral results in older adults and showed that the age-related decline in the learning of state transition structures can be mimicked in younger adults by inhibiting prefrontal cortex function using repetitive transcranial magnetic stimulation (rTMS) (Wittkuhn et al., 2018). Moreover, recent findings show that even when demands on the representation of the state transition structure are reduced, older adults show marked difficulties learning probabilistic transitions between states when compared to younger adults (Ruel et al., 2022). Consistent with the behavioral deficits, electroencephalography (EEG) data show blunted trial-by-trial neural responses following unexpected state transitions in older compared to younger adults whereas neural responses to reward feedback seem intact in the elderly. These results indicate that the older adults fail to realize changes in the task structure (as indicated by the lack of state prediction errors in the ERP), whereas they seem to be able to adjust behavior as a function of reward feedback.

Taken together, the current findings suggest that the observed age-related deficits in older adults in the representation and updating of state transition structures (state spaces) result from functional decline in the (orbito-)frontal cortex (OFC) and possibly hippocampus. While this neurobiological explanation is consistent with behavioral findings and computational accounts of

decision-making in older adults, there is a strong need for further research to support this hypothesis. We suggest that future work should first examine the OFC and HC independently in order to improve our understanding of the impact of aging on these regions and the effects this has on older adult's ability to engage in goal-directed learning. Based on the findings in younger adults, we predict that the hippocampus may be involved in helping the user represent transitions between states, while the OFC may play a role in helping the user apply and update state space representations. Methods such as MVPA and computational modeling may help develop our mechanistic understanding of the representational capacities of the HC and OFC.

### **Summary of the Theory**

In the diminished state space theory of human aging, we propose that age-related limitations in complex goal-directed behavior may result from an underlying deficit in the representation of state spaces of cognitive tasks (see Figure 1). We provide evidence from behavioral and cognitive neuroscience studies that support our theory, and we make first steps towards potential computational implementations of such deficits. As outlined in the introduction, in industrialized high-tech societies there is an increased need for complex learning and decision-making abilities. However, these abilities clearly decline with advancing age and the societal impact of this decline is potentiated given the demographic changes. To be able to counteract the implications of these age-related limitations, we need to understand the underlying cognitive, computational and neurobiological mechanisms. Here we propose that diminished state space representations might be one source of these deficits.

The diminished state space theory could be understood as an 'intermediate' level theory that provides an interface between lower-level deficits in basic cognitive processes and interactions with the environment. We think that this is the strength of the theory rather than a weakness. In the last decades, several theories have been proposed to explain cognitive aging in terms of cognitive primitives, such as working memory (e.g. (Kirasic et al., 1996; Park et al., 2003), inhibition (Lustig et al., 2007), speed of processing (Salthouse, 2000), or episodic memory (e.g., Shing et al., 2010). These theories have been successful in explaining age differences in cognition but each of them in a relatively limited domain, in specific tasks, or with the additional assumption that the corresponding cognitive process might contribute significantly to some higher-level deficit (e.g., complex decision-making, reasoning or cognitive control).



Instead of focusing on one of these primitives, we suggest that age-related changes in goal-directed cognitive behavior result from a deficit on this intermediate level, the representation of the state space of a cognitive task. This does not mean that the cognitive primitives do not play a role for explaining cognitive aging. Rather, we believe the opposite to be true. We explicitly assume that deficits in these cognitive primitives may contribute to impoverished state space representations and describe candidate mechanisms on the computational and neurobiological level.

### **Specific predictions of the theory, generalization and pathways for application**

**Specific predictions.** In the paragraph on the computational mechanisms, we suggest that diminished state space representations in older adults may result from the pruning branches of the decision tree, from the use of simplified task construals, or from the application of heuristic strategies such as successor representation learning. The pruning of branches of the decision tree is difficult to assess directly in experimental paradigms, but could be studied indirectly by testing memory for state information encoded during learning. Differences in subjective task representations (construals) have been investigated in a recent study by Ho et al. (2022) using a goal-directed maze navigation paradigm. Using these types of tasks in older adults should reveal that deficits in forward planning and navigation in the elderly result from over-simplified task construals. Finally, recent work has studied the use of successor representation learning as a cognitively less demanding alternative decision strategy (Momennejad et al., 2017). This approach could be used in older adults in combination with neural measures of offline replay (Momennejad et al., 2018). The hypothesis would be that diminished state spaces in older adults reflect the use of successor representation learning and are associated with differences in replay. In the paragraph on the computational mechanisms underlying deficits in state space representations, we also suggest that an inefficient updating of transition probabilities may be one source of impoverished state space representations in older adults. This could be tested empirically by comparing neural (EEG or MEG) correlates of state predictions errors (SPE's) of different magnitude in conditions in which SPE's can or cannot be used for learning. This would allow to dissociate age-related deficits in SPE signaling itself from an inappropriate updating of state predictions in older adults (see (Nassar et al., 2019)).

**Generalization to other cognitive domains.** One cognitive domain in which older adults show consistent deficits is cognitive control. An important aspect of control in the Stroop task or the task switching paradigm is to be able to accurately represent the task sets (defined as all stimulus–action mappings required to perform a task) and to keep the representation of task sets as separate from each other as possible (see (Musslick & Cohen, 2021)). In line with the diminished state space theory, task set representations diminish with aging, which, in turn, should lead to specific impairments in task switching in older adults. That is, based on the theory, we predict that a diminished representation of task sets should affect performance on switch and repeat trials during task switching (general or global switch costs) whereas it should leave the actual process of switching unaffected (specific or local switch costs). This is consistent with what is observed empirically in the aging and task switching literature (for a meta-analysis see (Wasylyshyn et al., 2011)). Future research could try to test the predictions of the diminished state space theory in task-switching paradigms by further increasing the load on the representation of the state space (e.g., by increasing the number of tasks) or by manipulating the local transition probabilities of the task sets.

The diminished state space theory could also be applied to explain age-related deficits in reversal learning or performance on tasks such as the Wisconsin Card Sort Test. Work by Hämmerer et al. (2019), suggests that older adults tend to overestimate the likelihood of reversals during learning and they interpret their findings as a deficit in building up accurate representations of optimal choice behavior (Hämmerer et al., 2019). Viewed through the lens of the diminished state space theory, it could be argued that the underlying deficit might be in establishing differentiated state representations. Recent data in rats suggests that state representations during reversal learning can be decoded from activity of orbitofrontal neurons (Bartolo & Averbeck, 2020; Stalnaker et al., 2021). In humans, the work by Schuck and Niv (2019) showed similar representations in the OFC, which were associated with task performance in a variant of a reversal learning task. Based on the work by Schuck & Niv (2019), the most straightforward way to test the diminished state space theory in task switching and reversal learning tasks would be to try and decode task sets and attentional states from fMRI data during performance of these tasks. Accordingly, these state representations should be less differentiated

in older adults and the representational deficits should predict learning and task switching performance in the elderly.

**Applications.** To exemplify potential applications of the theory we come back to the initial example of an older adult trying to purchase a train ticket. Apart from the in-person ticket counter there are two ways of purchasing a train ticket: At a ticket vending machine or using a computer app. Both options are characterized by complex state spaces that are not directly observable and have been shown to be a major challenge for older adults (Schreder et al., 2012; Sengpiel, 2016). One important feature that distinguishes vending machines from apps is the degree to which the device can learn about the preferences and cognitive abilities of their users. This in turn determines design recommendations. One obvious way to support model-based decision-making in older adults while using these devices would be to reduce overall representational load by simplifying the state space (the number of available states and actions). Another way could be to foreshadow transitions to subsequent states (menus) or to provide an a priori map of the state space. Thereby one could offload the computational cost of having to internally represent and update state transitions. Lindenberger & Mayr (2014) suggest that environmental cues (such as on a menu of a ticket vending machine) should be compatible (functionally related to the goal of the agent) and distinctive (they should not co-activate competing actions). Compatibility in multistage decision making would mean that at the start state decision trajectories that are in line with the predicted preferences of the agent would be made more available. Furthermore, the state space could be adapted such that it matches the task construals or simplifying decision strategies of the user. Distinctiveness could mean that state transition structures need to be unambiguous and predictable. However, as outlined above, surprising outcomes (state transitions or rewards) may trigger learning. Thus, it could be useful to consider ways to induce surprising transitions (e.g., to rewarding end states) in order to support learning in the agent. The potential applications of the diminished state space theory outlined above center on reducing load on state space representations in older adults. However, there are two interrelated trade-offs to be considered: (1) in order to maximize learning in older adults (and other age groups) it may be important to account for individual cost-benefit evaluations and to adjust the decision environment accordingly (see Devine et al., 2021; Ruel, Devine, et al., 2021). (2) An over-reliance on environmental support (such as the adaptive

algorithms outlined above) may diminish cognitive resources in older adults (see Lindenberger & Mayr, 2014). Finally, there are ethical and data security concerns that need to be considered. After all, the interest of the transportation company may not align with that of the customer and the algorithms underlying such technology should remain transparent to its users.

### **Relationship to other theories**

The diminished state space theory ties in with suggestions that older adults rely more on external information across several psychological domains. Lindenberger and Mayr (2014) propose that this greater reliance on environmental support might be (mal-) adaptive adjustment to deficits in the ability to trigger and maintain cognitive representations (Lindenberger & Mayr, 2014). Here we provide a first explanatory framework for what these representations are and how they might be implemented computationally and on a neurobiological level. Future work should try to capture differences in the distinctiveness of state space representations across groups (e.g., through MVPA) and use this information to predict shifts in decision strategies as well as a greater reliance on environmental control. The current theory can be seen as being foreshadowed by past work on working with memory (Moscovitch, 1994) in which the author suggests that memory performance does not reflect the operation of a single system but depends on the interaction of processes within a network of interrelated components including sensory areas, basal ganglia, hippocampus and the prefrontal cortex. The diminished state space model also resonates with the neuronal gain control theory of human aging, which proposes that less differentiated cognitive representations in older adults may result from deficient dopaminergic input to cortical areas (Li et al., 2001). However, so far there is no direct evidence for an association between age-related decline in the dopamine system and deficits in state space representations in the OFC and hippocampus. This should be established in the future. Our theory also seems compatible with the proposal that age-related changes in the locus coeruleus (LC) norepinephrine system may underlie cognitive decline in older adults (Mather & Harley, 2016). Mather and colleagues suggest that the phasic norepinephrine (NE) signals from the LC may shift cognitive representations in accordance with arousal levels (Mather et al., 2018). Specifically, norepinephrine may modulate cortical excitation and inhibition, which in turn, guides selective processing of information during learning. Therefore, age-related degeneration of the LC system may result in impairments in selective attention and downstream effects on model-based

learning. Although some research shows that older adults have a less reliable attentional filter (Dahl et al., 2020), future research should consider this research question using indicators of LC function (such as pupillometry and EEG signals) as well as structural (MRI) measures of LC integrity to predict age-related changes state space representations. Finally, we would like to conjecture that deficits in state space representations may not necessarily imply suboptimal learning and decision behavior in older adults. From a resource-rational perspective (Lieder & Griffiths, 2020), it could well be that shifts in behavioral strategies with age reflect a boundedly optimal adaptation to internal constraints (diminished state space representations) in the elderly (Devine et al., 2021; Ruel, Devine, et al., 2021).

## References

- Barbey, A. K., Koenigs, M., & Grafman, J. (2011). Orbitofrontal Contributions to Human Working Memory. *Cerebral Cortex*, 21, 789–795.
- Bartolo, R., & Averbeck, B. B. (2020). Prefrontal cortex predicts state switches during reversal learning. *Neuron*, 106, 1044-1054.
- Bolenz, F., Kool, W., Reiter, A. M. F., & Eppinger, B. (2019). Meta-control of decision-making strategies in human aging. *eLIFE*, 8, 1-24.
- Bornstein, A. M., & Daw, N. D. (2013). Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLOS Computational Biology*, 9.
- Bornstein, A. M., & N.D., D. (2013). Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLOS Computational Biology*, 9.
- Bradfield, L. A., Dezfouli, A., van Holstein, M., Chieng, B., & Balleine, B. W. (2015). Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron*, 88, 1286- 1280.
- Chan, S. C. Y., Schuck, N. W., Lopatina, N., Schoenbaum, G., & Niv, Y. (2021). Orbitofrontal cortex and learning predictions of state transitions. *Behavioral Neuroscience*, 135, 487-497.
- Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Duzel, E., & Dolan, R. J. (2013). Dopamine restores reward prediction errors in old age. *Nature Neuroscience*, 16, 648-653.
- Czaja, S., Charness, J., Fisk, N., Hertzog, A., Nair, S. C., Rogers, W., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (create). *Psychology and Aging*, 21, 333-352.
- Dahl, M. J., Mather, M., Sander, M. C., & Werkle-Bergner, M. (2020). Noradrenergic responsiveness supports selective attention across the adult Lifespan. *Journal of Neuroscience*, 40, 4372-4390.
- 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, 1204-1215.
- Deuker, L., Bellmund, J. L. S., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *eLife*, 5, 1-26.
- Devine, S., Neumann, C., Levari, D., & Eppinger, B. (2022). Human ageing is associated with more rigid concept spaces. *Psychonomic Bulletin and Review*.

- Devine, S., Neumann, C., Otto, A. R., Bolenz, F., Reiter, A. F., & Eppinger, B. (2021). Seizing the Opportunity: Lifespan Differences in the Effects of the Opportunity Cost of Time on Cognitive Control. *Cognition*, 216.
- Doeller, C. F., King, J. A., & Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Sciences*, 105, 5915- 5920.
- D'Acromont, M., & Bossaerts, P. (2016). Neural mechanisms behind Its identification of leptokurtic noise and adaptive behavioral response. *Cerebral Cortex*, 26, 1818-1830.
- Eppinger, B., Haemmerer, 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.
- Eppinger, B., Heekeren, H. R., & Li, S. C. (2015). Age-related prefrontal impairments implicate deficient prediction of future reward. *Neurobiology of Aging*, 36, 2380-2390.
- Eppinger, B., Schuck, N. W., Nystrom, L. E., & Cohen, J. D. (2013). Reduced striatal responses to reward prediction errors in older compared to younger adults. *Journal of Neuroscience*, 33, 9905-9912.
- Eppinger, B., Walter, M., Heekeren, H. R., & Li, S.-C. (2013a). Of goals and habits: Age-related and individual differences in goal-directed decision-making [Original Research]. *Frontiers in Neuroscience*, 7.
- Gershman, S. J., & N.D., D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, 68, 101-128.
- Gläscher, J., Daw, N. D., 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, 585-595.
- Hämmerer, D., Schwartenbeck, P., Gallagher, M., Fitzgerald, T. H. B., DüzelE, & Dolan, R. J. (2018). Older adults fail to form stable task representations during model-based reversal inference. *Neurobiology of Aging*, 74, 90-100.
- Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature neuroscience reviews*, 5, 87-96.
- Ho, M. K., Abel, D., Correa, C. G., Littman , M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan Nature, 606, 129-136.

- Huys, Q., Eshel, N., O’Nions, E., Sheridan, L., P., D., & Roiser, J. P. (2012). Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLOS Computational Biology*, 8.
- Huys, Q., Lally, N., Faulkner, P., Seifritz, E., Gershman, S. J., P., D., & Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 112, 3098- 3103.
- Johnson, A., & Reddish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27, 12176-12189.
- Johnson, E. L., Chang , W. K., Dewar , C. D., Sorensen , D., Lin , J. J., Solbakk , A.-K., . . . Knight , R. T. (2022). Orbitofrontal cortex governs working memory for temporal order. *Current Biology*, 32, R410-R411.
- Kaplan, R., KIng, J., Koster, R., Penny, W. D., Burgess, N., & Friston, K. (2017). The neural representation of prospective choice during spatial planning and decisions. *PloS Biology*, 15.
- Keramati, M., Smittenaar, P., Dolan, R. J., & P., D. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proceedings of the National Academy of Sciences*, 113, 12868-12873.
- Kirasic, K. C., Allen, G. L., Dobson, S. H., & Binder, K. S. (1996). Aging, cognitive resources, and declarative learning. *Psychology and Aging*, 11, 658-670.
- Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, 6, 691-702.
- Lengyel, M., & P., D. (2007). Hippocampal contributions to control: The third way *Advances in Neural Information Processing Systems*, Vancouver.
- Lerner, I., Sojitra, R., & Gluck, M. A. (2018). How age affects reinforcement learning. *Aging*, 10, 3630- 3631.
- Leung, R., Tang, C., Haddad, S., McGrenere, J., & Graf, P. (2012). How older adults learn to use mobile devices: survey and field investigations. *ACM Transactions on Accessible Computing*, 4, 1-33.
- Li, S.-C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: From neuromodulation to representation. *Trends in Cognitive Sciences*, 5(11), 479-486.
- Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43.



- Lindenberger, U., & Mayr, U. (2014). Cognitive aging: is there a dark side to environmental support? *Trends in Cognitive Sciences*, 18, 7-15.
- Liu, Y., Dolan, R. J., Kurth-Nelson, Z., & Behrens, T. E. (2019). Human Replay Spontaneously Reorganizes Experience. *Cell*, 178, 640-652.
- Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D., & Dolan, R. J. (2021). Experience replay is associated with efficient nonlocal learning. *Science*, 372, 1-9.
- Lustig, C., Hasher, L., & Zacks, R. (2007). Inhibitory Deficit Theory: Recent Developments in a "New View". In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 145–162). American Psychological Association.
- Malik, K. (2014). Human development report 2014: Sustaining human progress: Reducing vulnerabilities and building resilience. United Nations Development Programme.
- Markant, D. B. (2020). Active transitive inference: When learner control facilitates integrative encoding. *Cognition*, 200.
- Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2018). Norepinephrine ignites local hot spots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, 39.
- Mather, M., & Harley, C. W. (2016). The locus coeruleus: Essential for maintaining cognitive function and the aging brain. *Trends in Cognitive Sciences*, 20, 214-226.
- Momennejad, I., Otto, A. R., N.D., D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. *eLIFE*, 7, 1-25.
- Momennejad, I., Russek, E. M., Cheon, J. H., Botvinick, M. M., N.D., D., & Gershman, S. J. (2017). The successor representation in human reinforcement learning. *Nature Human Behavior*, 1, 680-692.
- Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 269–310). The MIT Press.
- Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control. *Trends in Cognitive Sciences*, 25, 757-775.
- Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. *eLife*, 8.

- Nassar, M. R., Bruckner, R., Gold, J. I., Li, S. C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, 7, 1-13.
- Nogueira, R., Abolafia, J. M., Drugowitsch, J., Balaguer-Ballester, E., Sanchez-Vives, M. V., & MorenoBote, R. (2016). Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. *Nature Communications*, 8.
- O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a cognitive map*. Oxford University Press.
- Olafsdottir, H. F., Bush, D., & Barry, C. (2018). The role of hippocampal replay in memory and planning. *Current Biology*, 28, 37-50.
- Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, 110, 20941-20946.
- Otto, A. R., Skatova, A., Madlon-Kay, S., & N.D., D. (2015). Cognitive control predicts use of model-based reinforcement-learning. *Journal of Cognitive Neuroscience*, 27, 319-333.
- Park, D. C., Welsh, R. C., Marshuetz, C., Gutchess, A. H., Mikels, J., Polk, T. A., . . . Taylor, S. F. (2003). Working memory for complex scenes: age differences in frontal and hippocampal activations. *Journal of Cognitive Neuroscience*, 15, 1122-1134.
- Rajah, M. N., & d'Esposito, M. (2005). Region-specific changes in prefrontal function with age: a review of PET and fMRI studies on working and episodic memory. *Brain*, 128, 1964-1983.
- 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*.
- Resnick, S. M., Lamar, M., & Driscoll, I. (2007). Vulnerability of the orbitofrontal cortex to age-associated structural and functional brain changes. *Annals of the New York Academy of Sciences*, 1121, 562- 575.
- Rmus, M., Ritz, H., Hunter, L. E., Bornstein, A. M., & Shenhav, A. (2019). Individual differences in model-based planning are linked to the ability to infer latent structure. In *bioRxiv*.
- Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10, 184-194.
- Ruel, A., Bolenz, F., Li, S.-C., Fischer, A. G., & Eppinger, B. (2022). Neural evidence for age-related deficits in the representation of state spaces. *Cerebral Cortex*.

- Ruel, A., Devine, S., & Eppinger, B. (2021). Resource-rational approach to meta-control problems across the lifespan. *WIREs Cognitive Science*, 12. <https://doi.org/doi:10.1002/wcs.1556>
- Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, 54, 35-54.
- Schad, D. J., Juenger, E., Sebold, M., Garbusow, M., Bernhardt, N., Javadi, A. H., . . . Huys, Q. (2014). Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Frontiers in Psychology*, 5, 1-10.
- Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, 10, 885-892.
- Schreder, G., Siebenhandl, K., Mayr, E., & Smuc, M. (2012). The ticket machine challenge: social inclusion by barrier-free ticket vending machines. In E. F. Loos, L. Haddon, & M.-M. E (Eds.), *Generational Use of New Media* (pp. 129-148). Ashgate Publishing.
- Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91, 1402-1412.
- Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, 364, 1-9.
- Sengpiel, M. (2016). Teach or Design? How Older Adults' Use of Ticket Vending Machines Could Be More Effective. *ACM Transactions on Accessible Computing*, 9, 2-27.
- Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Li, S. C., Müller, V., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews*, 34, 1080-1091.
- Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*, 60, 197-200.
- Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A metaanalysis. *Psychology and Aging*, 10, 527-539.
- Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, 20, 1643-1652.

- Stalnaker, T. A., Raheja, N., & Schoenbaum, G. (2021). Orbitofrontal State Representations Are Related to Choice Adaptations and Reward Predictions. *Journal of Neuroscience*, 41, 1941–1951.
- Stawarczyk, D., & Wahlheim, C. N. E., J.A.Snyder, A.Z.Zacks, J.N. (2020). Aging and the encoding of changes in events: The role of neural activity pattern reinstatement. *Proceedings of the National Academy of Sciences*, 117, 29346–29353.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. MIT Press. T
- Takahashi, Y. K., Roesch, M. R., Wilson, R. C., Toreson, K., O'Donnell, P., Niv, Y., & Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neuroscience*, 14(1590-1599).
- Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A map for social navigation in the human brain. *Neuron*, 87, 231-243.
- United Nations, D. o. E. a. S. A. P. D. (2019). *World population prospects 2019: Highlights*.
- Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., . . . N.D., D. (2019). Hippocampal contributions to model-based planning and spatial memory. *Neuron*, 102, 683-693.
- Wallis, J. D. (2007). Orbitofrontal cortex and its contribution to decision making. *Annual Review of Neuroscience*, 30, 30-56.
- Wang, F., Schoenbaum, G., & Kahnt, T. (2020). Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PloS Biology*, 18.
- Wasylshyn, C., Verhaeghen, P., & Sliwinski, M. J. (2011). Aging and task switching: a meta-analysis. *Psychology and Aging*, 26, 15-20.
- West, R. L. (1996). An application of prefrontal cortex functioning theory to cognitive aging. *Psychological Bulletin*, 120, 272-292.
- Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, 81, 267-279.
- Wimmer, E., & Büchel, C. (2019). Learning of distant state predictions by the orbitofrontal cortex in humans. *Nature Communications*, 10.

Wittkuhn, L., Eppinger, B., Bartsch, L., Thurm, F., Korb, F. M., & Li, S.-C. (2018). Repetitive transcranial magnetic stimulation over dorsolateral prefrontal cortex modulates value-based learning during sequential decision-making. *Neuroimage*, 167, 384-395.

## CHAPTER 4: ADULT AGE DIFFERENCES IN THE UPDATING OF STATE SPACES

Alexa Ruel<sup>1,3</sup> • Ben Eppinger<sup>2,3</sup>

<sup>1</sup> Department of Psychology, Concordia University, Montreal, Canada

<sup>2</sup> Department of Psychology, University of Greifswald, Greifswald, Germany

<sup>3</sup> PERFORM centre, Concordia University, Montreal, Canada

## Abstract

An increasing number of empirical studies suggest that older adults experience difficulties engaging in goal-directed decision-making due to their deficits in representing an abstract model of the task structure (i.e., state space), including the states, the actions that can be taken at each state and the transition structure that connects the states. From a computational perspective, one possible explanation for older adults' deficits in representing state spaces is their reduced efficiency in updating their internal model of the state space of the task. In the current study, we used a novel paradigm in which younger and older adults made decisions in two conditions (reversal learning vs. oddball) that are identical with respect to their perceptual input but differ in terms of the relationship between surprise and the need to update the internal representation of the task. We were therefore able to examine if older adults differ from younger adults in their ability to discriminate between surprising outcomes that can be safely ignored, from those that dictate the need to update their internal model. Behavioral results suggest that older adults showed greater impairments in the reversal learning condition as compared to the oddball condition than younger adults did. Our ERP analyses revealed that in contrast to the younger adults, older adults demonstrated reduced updating signals and showed no significant relationship between these signals and behavior in the reversal learning condition. We therefore propose that older adults' behavioral difficulties in the reversal learning condition may arise from a deficit in discriminating between common and rare outcomes in this condition. In line with the diminished state space theory, we find preliminary evidence that older adults' deficits in creating an abstract model of the task, may in part be due to their deficit updating their internal model.

## Introduction

The ability to make goal-directed decisions that rely on mental models of the environment has been shown to decline during healthy aging (Eppinger et al., 2013; Bolenz et al., 2019; Ruel et al., 2022). One recent theory suggests that difficulties in goal-directed decision-making arise from an underlying deficit in older adults' ability to learn and represent the state spaces of cognitive tasks (Eppinger et al., 2023). In this theory, a state space is defined as an abstract representation of the structure of the task, including the states (stimuli), the available actions and the transition structure that connects the different states. From a computational perspective, older adults' representational deficits are hypothesized to stem from two, not mutually exclusive, sources: 1) an inaccurate representation of states, actions, or transition probabilities or 2) an inefficient updating of these transition probabilities. Thus far, behavioral results (Bolenz, et al., 2019; Eppinger et al., 2015; Hämmerer et al., 2019; Ruel et al., 2022) have been incapable of disambiguating between these two possibilities. Therefore, it remains unclear if older adults experience deficits in representing and/or a reduced updating of their state space during decision-making.

Recent work has emphasized the important role that learning processes play when making decisions in complex and dynamically changing environments (e.g. D'Acremont & Bossaerts, 2016; Nassar et al., 2019). Namely, in volatile environments, where changes from one state to another may represent important information about a change in the configuration of the task, agents must employ a set of cognitive rules to identify which outcomes are important and which ones they can disregard. From a computational standpoint, recent work has demonstrated that learning principles based in temporal-difference learning can be leveraged to explain how humans (and other biological agents) navigate such stochastic environments (Bolenz et al., 2019; Gläscher et al., 2010). From this perspective, agents must dynamically update their representation of a task environment based on the states they encounter and the transitions between them. Therefore, proper encoding of the transitions between states is a critical component to forming a correct representation of the task environment. That is, the degree to which state outcomes ought to be encoded depends, critically, on the statistical properties of the environment in which they are situated (D'Acremont & Bossaerts, 2016; Nassar et al., 2019). For example, in changing environments, surprising outcomes dictate the need to update the internal



model of the environment. In contrast, in stable environments with occasional outliers (i.e., oddballs), a surprising outcome does not require that the agent update their internal model (Nassar et al., 2019).

Younger adults have been shown to be quite remarkable at assessing when a surprising outcome requires them to update their internal model and consequently to either emphasize or deemphasize surprising information (Cheadle et al., 2014; d’Acremont and Bossaerts, 2016; Nassar et al., 2010; Nassar et al., 2019; Summerfield & Tsetsos, 2015). In contrast, older adults may experience greater difficulties identifying outcomes that signal a need to update their representations (Samanez-Larkin et al., 2014; Nassar et al., 2016; Weiler et al., 2008). For example, several studies have demonstrated that older adults show impaired reversal learning (Hämmerer et al., 2019; Weiler et al., 2008) as well as a reduced neural representation of prediction errors during feedback-driven reinforcement learning (Eppinger et al., 2013; Samanez-Larkin et al., 2014). As prediction errors provide feedback signals that guide subsequent updates to the model’s internal representation (Gläscher et al., 2010), age-related deficits in computing prediction errors could have a significant impact on internal model updating. In line with this, Nassar and colleagues (2016) have shown that age-related deficits in the flexible adjustment of transition learning rates might lead to an insufficient (i.e., rigid) representation of the environment. Taken together, it seems that while younger adults have the ability to efficiently update their internal representation of the task environment based on feedback obtained during the task, older adults may experience difficulties identifying when an outcome should be learned from, which could contribute to their deficits in representing state spaces.

One method employed by past research to examine how adjustments in individuals learn and consequently update their internal model in changing environments is electroencephalography (EEG) (see Jepma et al., 2016; Eppinger et al., 2017). Specifically, prior work using EEG has revealed that the P3b component of the event-related potential tracks learning and changes in learning rates across statistical contexts (Nassar et al., 2019; Wyart, et al., 2012). Specifically, the context updating theory (Donchin, 1981) suggests that the P3b, a positive stimulus-locked component, seems to reflect brain activity underlying a revision of the mental representation (Donchin, 1981; Polich, 2007). That is, even after controlling for the degree of surprise, this positive component relates to learning (Fischer & Ullsperger, 2013;

Jepma et al., 2018). In contrast, the P3a, originating from stimulus-driven frontal attention mechanisms during stimulus processing is thought to be unrelated to updating processes (Polich, 2007; Squire, 1975). Rather, the P3a has been shown to be generated when a stimulus is surprising given the current task representation.

In the current study, we examined if older adults experience difficulties in updating their internal model in changing environments by designing a simple paradigm which would allow us to dissociate updating and surprise. Specifically, younger and older adults completed a goal-directed decision-making task in an oddball condition as well as a reversal learning condition. In the reversal learning condition, surprising outcomes indicate that the association between actions and outcomes have reversed, therefore dictating the need for the participant to update their internal model. In contrast, surprising outcomes dictate no change in the environment, and therefore no updating is required. In line with the context-updating theory, we predicted that reversal trials, but not oddball trials would lead to an increase in participants' learning rates resulting in a larger P3b component, as well as a behavioral adjustment on the immediately subsequent trial. Based on the hypothesis that older adults may have deficits in their ability to identify surprising outcomes and consequently update their internal model (Eppinger et al., 2013; Nassar et al., 2016; Samanez-Larkin et al., 2014), we predicted that older adults would show a reduced P3b component in the reversal learning condition in comparison to younger adults.

## **Method**

### **Participants**

Thirty-four healthy younger adults and 41 healthy older adults participated in the current study. Younger adults were recruited via the participant pool at Concordia University and through online advertisements. Older adults were recruited via Concordia University's PERFORM center's mailing list. We excluded any participant whose data was missing 20% of trials or more and who performed at chance (50% correct responses) (younger adults:  $n=3$ , older adults:  $n=12$ ). For the ERP analyses, all participants had to have at least 28 viable surprise trials in each condition. The final sample consisted of 30 younger adults (18-32 years,  $M_{age} = 23.16$  years,  $SD = 3.11$  years, 12 males) and 29 older adults (60-82 years,  $M_{age} = 70.48$  years,  $SD =$

5.16years, 13 males). All participants gave informed written consent before participating. The ethics committees of Concordia University approved the study. Participants received a minimum payment of \$20.00 CAD and a bonus of up to \$12.00CAD based on their performance on the task. Eligible younger adults who registered to participate in the study through the Concordia Participant Pool could receive 2 credits instead of \$20.00 CAD.

## **Stimuli**

Stimuli consisted of 2 differently colored houses and 2 differently colored buses. Original images taken from a royalty-free image website (vecteezy.com), modified in the free software Inkscape (inkscape.org) and made equally luminescent using the independently developed MATLAB SHINE color toolbox ([https://github.com/RodDalBen/SHINE\\_color](https://github.com/RodDalBen/SHINE_color)). The resulting stimuli consisted of a blue bus and a red bus, as well as two visually distinct and discriminable houses (see Fig 1). All stimuli were presented on a 23.6-inch VIEWPixx/EEG computer screen using Psychopy3 software (Peirce & MacAskill, 2018).

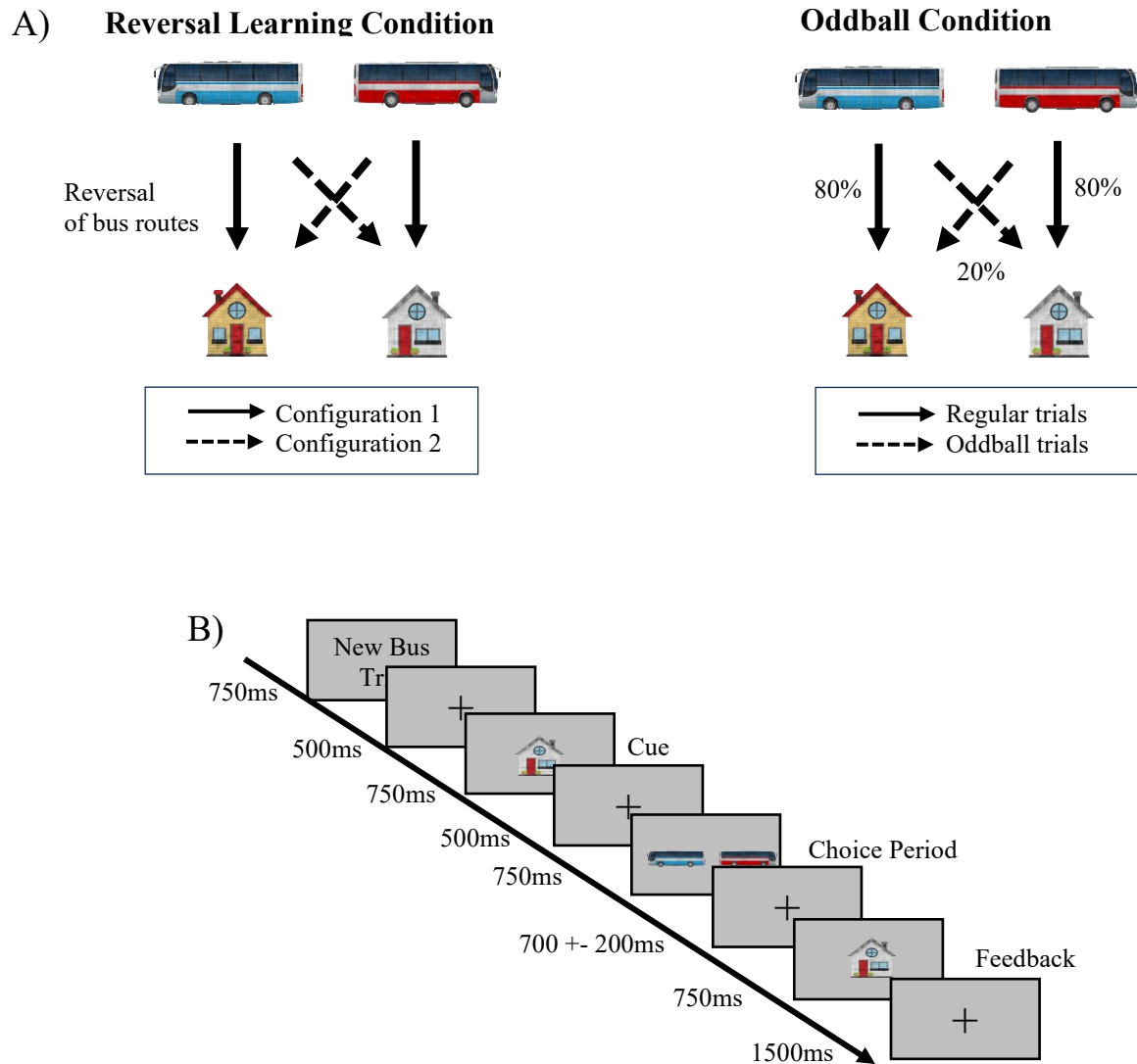
## **Task**

A novel task, with two conditions (oddball and reversal learning) was designed for the present study. The paradigm was programmed in Psychopy (Version 3; Pierce et al., 2019).

During the learning phase, participants had to learn the association between houses and buses by trial and error. On each trial, participants were first shown a cue, which consisted of one of the two houses, presented for 750ms. Following this, participants had to make a choice between the two buses (red and blue) by using the “A” and “L” keys on a computer keyboard. Buses were randomly counterbalanced to appear on the left or right side of the screen for 750ms. After this choice period, participants were presented with a fixation cross with a mean presentation time of 700ms and a standard deviation of 200ms. This variable interstimulus interval (ISI) reduced the likelihood that participants would anticipate the feedback presentation. Finally, during the feedback stage of the task, participants were presented with the house their chosen bus ‘traveled’ to (i.e., the house that bus is associated with). Participants were not presented with reward feedback per se. Rather, they had to infer the accuracy of their choice

based on the feedback stimuli they obtained. If participants made the correct bus choice, they would see the same (cued) house as feedback. If their choice was incorrect, they were shown the house their chosen bus travels to, which was not cued on that trial. If participants failed to make a choice between the two buses, they received “Too Slow” as feedback. Similarly, if participants pressed a key before the buses were presented, they received “Too Fast” as feedback. All feedback stimuli were presented for 750ms. Note that, for older adults, the task included an additional practice phase at the start of the task. This practice phase was added to help older participants learn the association between houses and buses before starting the practice phase with added time pressure. The practice phase ended once participants achieved 80% accuracy.

Once the practice phase was completed, participants started the experimental blocks. In the oddball condition, participants experienced one configuration of the task structure with the occasional oddball trial type on which the association between the houses and buses would switch. In the reversal learning condition, participants experienced one task configuration, until an unexpected outcome indicated that the association between houses and buses had reversed. Oddball trials and reversal trials (surprising trials) occurred on 20% of trials with a mean of 4 trials between oddball or reversal trial types (herein referred to as rare trial types) with a minimum of 2 trials and a maximum of 7 trials between rare trials. In both conditions, participants were not instructed on the occurrence of surprising outcomes, but only in the reversal learning condition did a surprising outcome dictate the need to update their internal model. Before the start of each block, participants were presented with instructions that explained its statistical context (i.e., the oddball or reversal learning context). For oddball blocks, participants were told that due to construction, sometimes the buses will swap routes, traveling to the opposite house for a single trial. Instructions for reversal learning blocks were similar, but participants were instead told that buses would swap routes for several trials in a row instead of just one. For both conditions, participants were warned that they would not be instructed when the buses would swap their routes. Yet, their goal always remained the same: to get as many trials correct. For every correct answer, participants were compensated \$0.04 CAD.



**Figure 1.** **A)** Schematic representation of the relationship between houses and buses according to the task configuration, with the reversal condition illustrated on the left, and oddball condition on the right. **B)** Example trial in the bus task. In this example, the participant would have made the correct choice between the two buses during the choice period, and therefore obtains feedback that confirms that their choice was correct.

## Procedure

All participants provided written consent to participate in the present study. While the experimenter prepared the electroencephalogram (EEG), participants completed a demographic

questionnaire. All participants first completed the learning phase, in which they were instructed to learn the association between the buses and houses. The learning phase ended once participants reached an accuracy of 80% or after a maximum of 36 trials. Participants then completed 6 blocks of 60 trials each (3 blocks per condition) for a total of 360 trials (180 per condition). Blocks alternated between conditions within participants and block order was counterbalanced across participants.

## **Data Analysis**

Behavioral data were analysed using the lme4 package (Bates et al. 2014) for R (version 4.1.3). The EEG analyses were completed in MATLAB (Mathworks Inc. Natick, MA) using EEGLAB (Delorme & Makeig, 2004).

### *Behavioral Data*

To examine our prediction that the reversal learning condition should result in behavioral adjustments in the immediately subsequent trial, we examined each trial type separately. Note that rare trials refer to the oddball trial in the oddball condition and to the reversal trials in the reversal learning condition. The rare+1 trial types were defined as trials immediately following the rare trials in both conditions. All trials in which the task configuration remained unchanged are referred to as a common trial type. We therefore ran mixed effects regressions on reaction time and accuracy data, with factors condition (oddball vs. reversal learning) and age group (younger vs. older adults) for each trial type. All regressions were implemented using the lme4 package (Bates et al., 2014) in R (R Development Core Team 2010).

## EEG Recording and Analysis

### *Preprocessing*

EEG was recorded continuously from 64 active Ag/AgCl electrodes embedded in an elastic cap, using Brain Vision Recorder (Brain Products GmbH, Gilching, Germany). The electrodes were placed according to the international 10-10 system. During the recording, electrodes were referenced to the left mastoid and re-referenced offline to the average of the left and right mastoids. The EEG signal was first filtered using a band pass filter in the range of 0.01 and 100 Hz and was digitized with a sampling rate of 1000 Hz. The ground electrode was placed

at the center of the forehead at Fpz. Electrode impedances were kept below 25 kOhm during recording.

For statistical analysis, data were resampled to 500 Hz and filtered using a band-pass filter of 0.5-20Hz. Next, bad channels were linearly interpolated, data was epoched and artifacts such as large muscle movements and large blinks were rejected by visual inspection. Visual rejection therefore resulted in 4.05% of trials being rejected for younger adults and 4.34 % for older adults. Data were epoched (-1 to 2s) surrounding the feedback phase and run through an independent component analysis (ICA) for further artifact rejection. Blinks, lateral eye movements and muscle components were identified in the ICA output and removed from the data. Finally, any trials on which participants failed to make a choice were removed from both the behavioral and EEG data.

#### *ERP analysis*

The 3s epochs were epoched (-200 to 1000ms) around the feedback stimuli presentation, where 0 corresponds to the feedback stimuli onset. These epochs were baseline corrected by subtracting the average activity within the first 200ms pre-stimulus activity from the entire epoch length. To ensure all condition and trial type averages had the same number of trials, we used the same subset of common trials as used for behavioral analyses which resulted in an average of 30 trials per type (common, rare and rare+1) for analysis.

*Stimulus-locked ERPs at the feedback stage.* We analysed two ERPs at the feedback stage of the task: the P300a (P3a) and the P300b (P3b). Due to known age-related differences in the latency of these components (see Pontifex, Hillman & Polich, 2009), the exact times selected for younger and older adults differ slightly. Nevertheless, for both age groups, component latencies were chosen based on the peak latencies, around which we built a 100ms time window for the analysis of the P3a and a 200ms time window for the analysis of the P3b. The P3a was measured as the mean amplitude in a 290-390 ms time window after stimulus onset for younger adults and in a 300-400 ms time window for older adults. Finally, the P3b was measured as the mean amplitude in a 400-600 ms time window after onset for younger adults, and in a 450-650 ms time window for older adults.

Each component was analysed at the electrode that showed the largest effects based on inspection of the topographical plots (see figures 3 & 4). Specifically, we analysed the P3a at Cz and the P3b at Pz.

### *Correlation analysis*

We investigated the relationship between behavioral and electrophysiological measures for each age group by calculating Pearson correlation coefficients. To do so, we computed amplitude difference values for the P3a and P3b in each condition (P3adiff\_OB, P3adiff\_RV, P3bdiff\_OB, P3bdiff\_RV). In line with our prediction that participants would exhibit behavioral adjustments following rare trial type in the reversal learning condition only, we focused on participants' performance on rare+1 trial type in each condition (acc\_OB, acc\_RV). Note a Bonferroni correction for multiple comparisons was applied, adjusted p-values are reported in the results.

## **Results**

### ***Behavioral Results***

Overall, younger adults were more accurate ( $B = 0.048$ ,  $SE=0.015$ ,  $p= 0.002$ ) as well as faster ( $B=-0.044$ ,  $SE=0.016$ ,  $p= 0.009$ ) than older adults.

#### **Common trial type.**

*Reaction times.* Results of the mixed effects regression on common trial type reaction times revealed main effects of age ( $B= -0.065$ ,  $SE=0.019$ ,  $p=0.001$ ), as well as a significant interaction between condition and age group ( $B= -0.012$ ,  $SE=0.004$ ,  $p=0.002$ ). To analyse the interaction effect, we ran separate analyses for the two age groups. While younger adults showed no difference in reaction times across conditions ( $B=-0.001$ ,  $SE=0.006$ ,  $p=0.842$ ), older adults were significantly slower in the reversal learning condition ( $B=0.019$ ,  $SE=0.005$ ,  $p < 0.001$ ).

*Accuracy.* The same analysis on common trial type accuracy data revealed main effects of age ( $B=0.061$ ,  $SE=0.023$ ,  $p=0.017$ ) and condition ( $B=-0.095$ ,  $SE=0.010$ ,  $p < 0.001$ ) as well as an interaction between condition and age group ( $B= 0.049$ ,  $SE= 0.011$ ,  $p < 0.001$ ). This analysis revealed that older adults were less accurate than the younger adults in the reversal learning



condition ( $B=0.099$ ,  $SE=0.037$ ,  $p=0.009$ ). In contrast, we found no difference across age groups for the oddball condition ( $B=0.033$ ,  $SE=0.022$ ,  $p=0.140$ ).

Together, in the oddball condition, older adults were significantly slower but did not differ from younger adults in their accuracy. In contrast, in the reversal learning condition, in comparison to younger adults, older adults were slower and less accurate.

### **Rare trial type.**

*Reaction times.* Results of the mixed effects regression on rare trial type reaction times revealed a main effect of age group ( $B= -0.060$ ,  $SE = 0.019$ ,  $p = 0.002$ ) and an interaction between condition and age group ( $B= -0.007$ ,  $SE = 0.004$ ,  $p=0.038$ ). To analyse this interaction, we performed separate analyses for the two age groups. Similar to the common trial type results, younger adults showed no difference in reaction times across conditions ( $B=-0.004$ ,  $SE=0.006$ ,  $p=0.479$ ), while older adults were significantly slower in the reversal learning condition ( $B=0.011$ ,  $SE=0.005$ ,  $p=0.017$ ).

*Accuracy.* Results from our accuracy analysis revealed a main effect of condition ( $B= 0.089$ ,  $SE= 0.011$ ,  $p< 0.001$ ) showing that both age groups were more accurate in the reversal learning condition. We also found a trending interaction between condition and age group ( $B= -0.021$ ,  $SE=0.011$ ,  $p= 0.051$ ). Performing analyses separately by condition revealed no significant difference in accuracy between age groups in the oddball condition ( $B= -0.022$ ,  $SE=0.029$ ,  $p= 0.442$ ). Yet in the reversal learning condition, older adults were more accurate than younger adults ( $B=-0.065$ ,  $SE=0.035$ ,  $p=0.073$ ).

Taken together, the analysis of rare trial types reveals that both age groups were more accurate in the reversal condition.

### **Rare+1 trial type.**

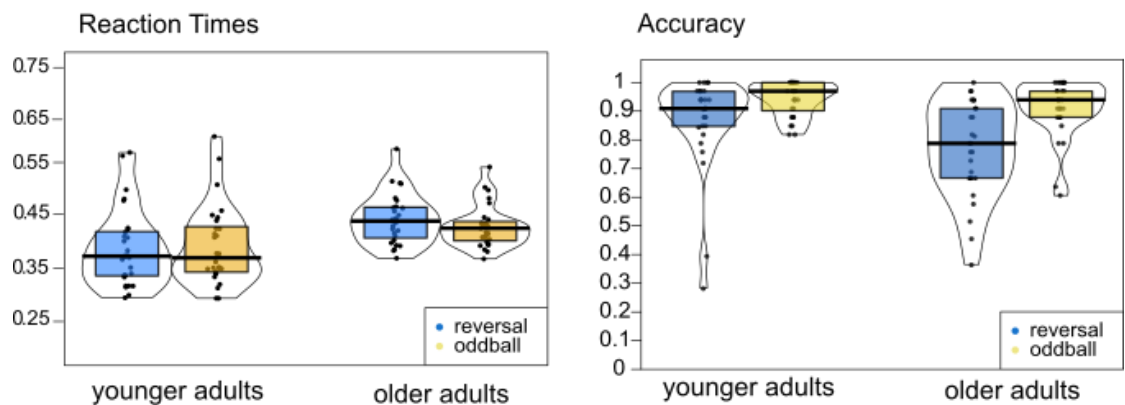
*Reaction times.* Results from the mixed effects regression on rare+1 trial type reaction times revealed a main effect of age ( $B= -0.065$ ,  $SE=0.019$ ,  $p= 0.001$ ) as well as a main effect of condition ( $B= -0.013$ ,  $SE=0.004$ ,  $p< 0.001$ ), demonstrating that older adults were significantly slower than younger adults, and that participants had faster reaction times in the reversal learning condition. However, contrary to our prediction, we found no significant condition by age interaction on reaction times ( $p=0.156$ ).

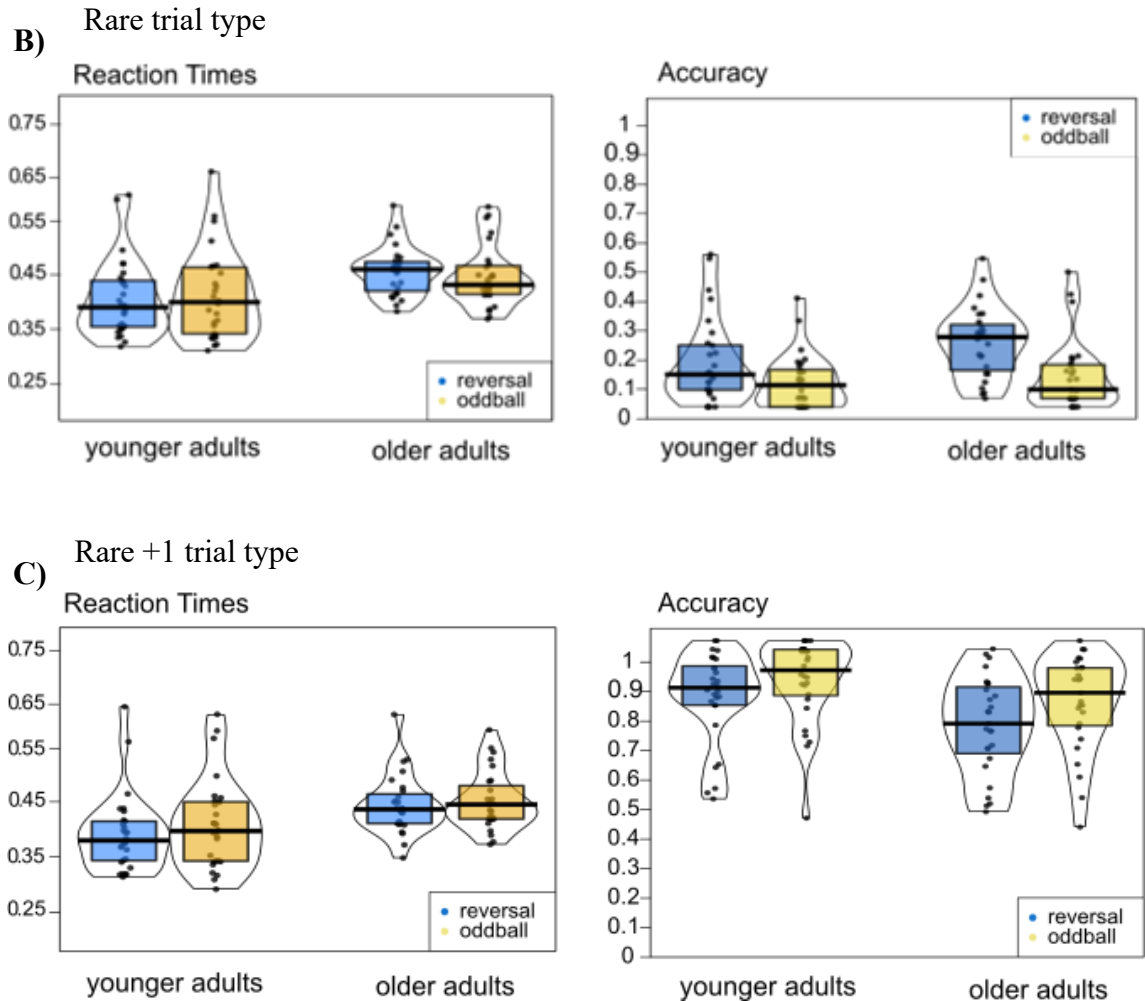
*Accuracy.* Results from the accuracy analysis revealed that older adults were less accurate than younger adults ( $B=0.083$ ,  $SE=0.031$ ,  $p=0.009$ ) and that both age groups were significantly less accurate in the reversal learning condition than in the oddball condition ( $B=-0.063$ ,  $SE=0.012$ ,  $p<0.001$ ). In line with the reaction time analysis, we found no condition by age group interaction.

Together, results of the analyses on rare+1 trial type suggest that both age groups were less accurate but faster in the reversal condition as compared to the oddball condition.

Overall, our behavioral results reveal that age differences were more pronounced in the reversal learning condition as compared to the oddball condition. In the reversal learning condition, older adults demonstrated slower and less accurate responses than younger adults did in this condition. In contrast, in the oddball condition, older adults were slower, but just as accurate as younger adults in both the common and rare trial types.

#### A) Common trial type





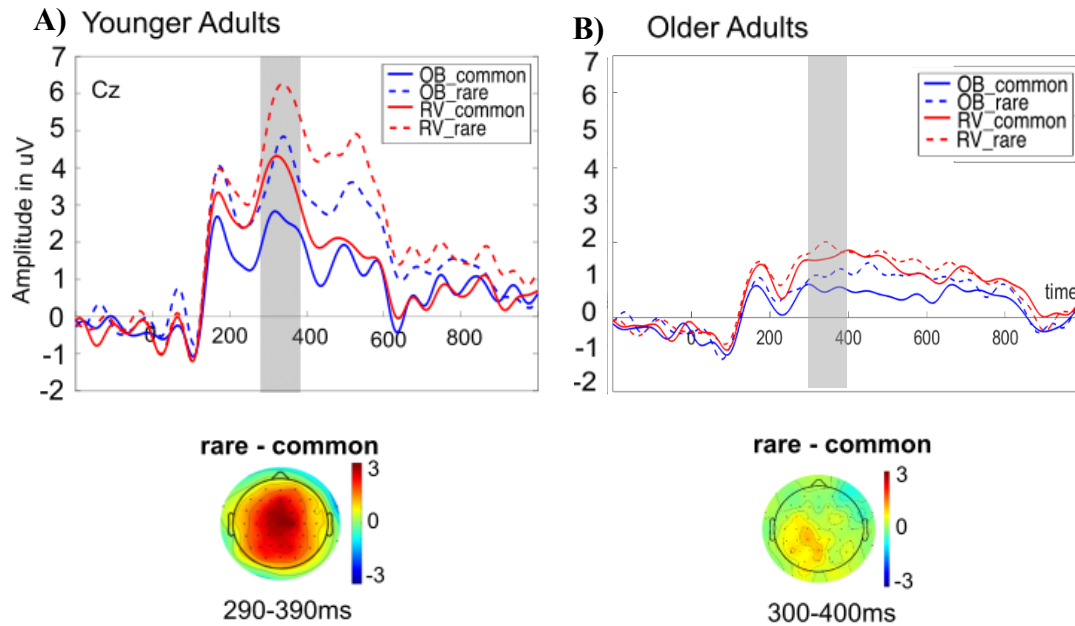
**Figure. 2.** Behavioral results (reaction times and accuracy) for both conditions across both age groups for the **A)** common trial type, **B)** rare trial type and **C)** rare +1 trial type. The reversal condition is reflected in blue bars, while the oddball condition is in yellow.

### ***ERP results***

#### **Stimulus-locked ERPs at the feedback stage**

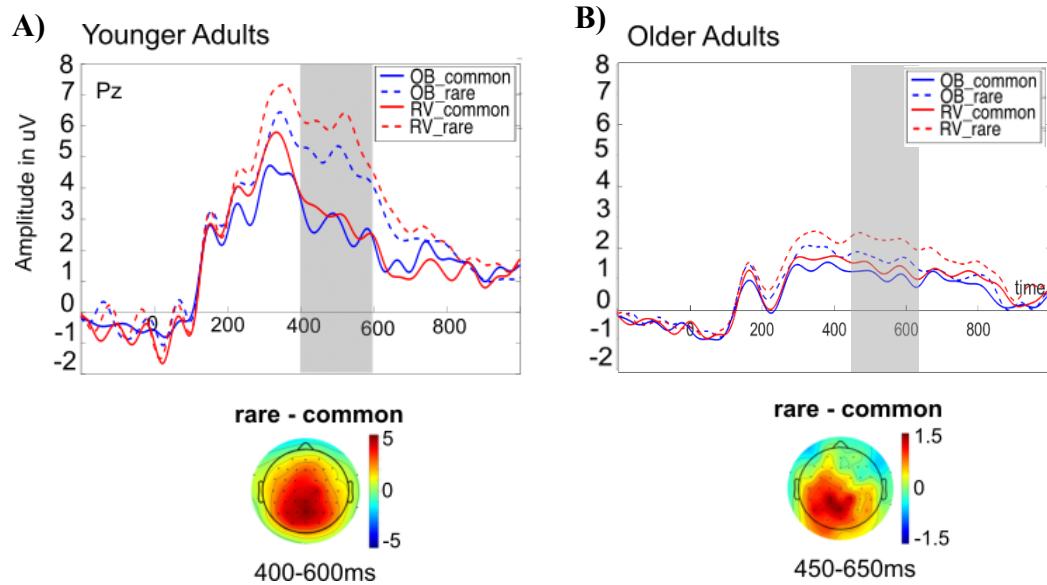
*P300a component.* The analysis of the P300a component at Cz revealed main effects of condition ( $B=0.546$ ,  $SE=0.215$ ,  $p=0.012$ ), trial type ( $B=-0.519$ ,  $SE=0.215$ ,  $p=0.017$ ) and age group ( $B=1.129$ ,  $SE=0.218$ ,  $p<0.001$ ). In contrast to our prediction, we found no significant condition by trial type interaction. However, we found a trending trial type by age interaction ( $B=-0.371$ ,  $SE=0.215$ ,  $p=0.08$ ). Follow-up analyses showed a larger P3a for rare compared to the

common trial type in younger adults ( $B=-0.890$ ,  $SE=0.218$ ,  $p<0.001$ ). In contrast, older adults show no difference in the P3a for the two trial types ( $B=-0.148$ ,  $SE=0.117$ ,  $p=0.210$ ).



**Figure 3. Top:** ERPs elicited by feedback stimuli at electrode Cz, displayed separately for the oddball condition (blue) and reversal learning condition (red), as well as the common trial type (solid lines), and rare trial type (dashed lines) for both younger **A)** and older adults **B)**. **Bottom:** topographical map displays of the difference between common and rare trials for younger and older adults.

*P300b.* The analysis of the P3b at Pz revealed main effects of trial type ( $B=-0.198$ ,  $SE=0.039$ ,  $p<0.001$ ) and age group ( $B=0.261$ ,  $SE=0.040$ ,  $p<0.001$ ), but no significant condition by trial type interaction. However, we did find a significant interaction between trial type and age group ( $B=-0.114$ ,  $SE=0.039$ ,  $p=0.004$ ). As displayed in figure 4, younger adults showed a larger P3b in response to rare as compared to common trial types ( $B=-0.312$ ,  $SE=0.036$ ,  $p<0.001$ ) than older adults did ( $B=-0.083$ ,  $SE=0.021$ ,  $p<0.001$ ).



**Figure 4.** Top: ERPs elicited by feedback stimuli at electrode Pz, displayed separately for the oddball condition (blue) and reversal learning condition (red), as well as the common trial type (solid lines), and rare trial type (dashed lines) for both younger **A**) and older adults **B**). Bottom: topographical map displays of the difference between common and rare trial types for younger and older adults.

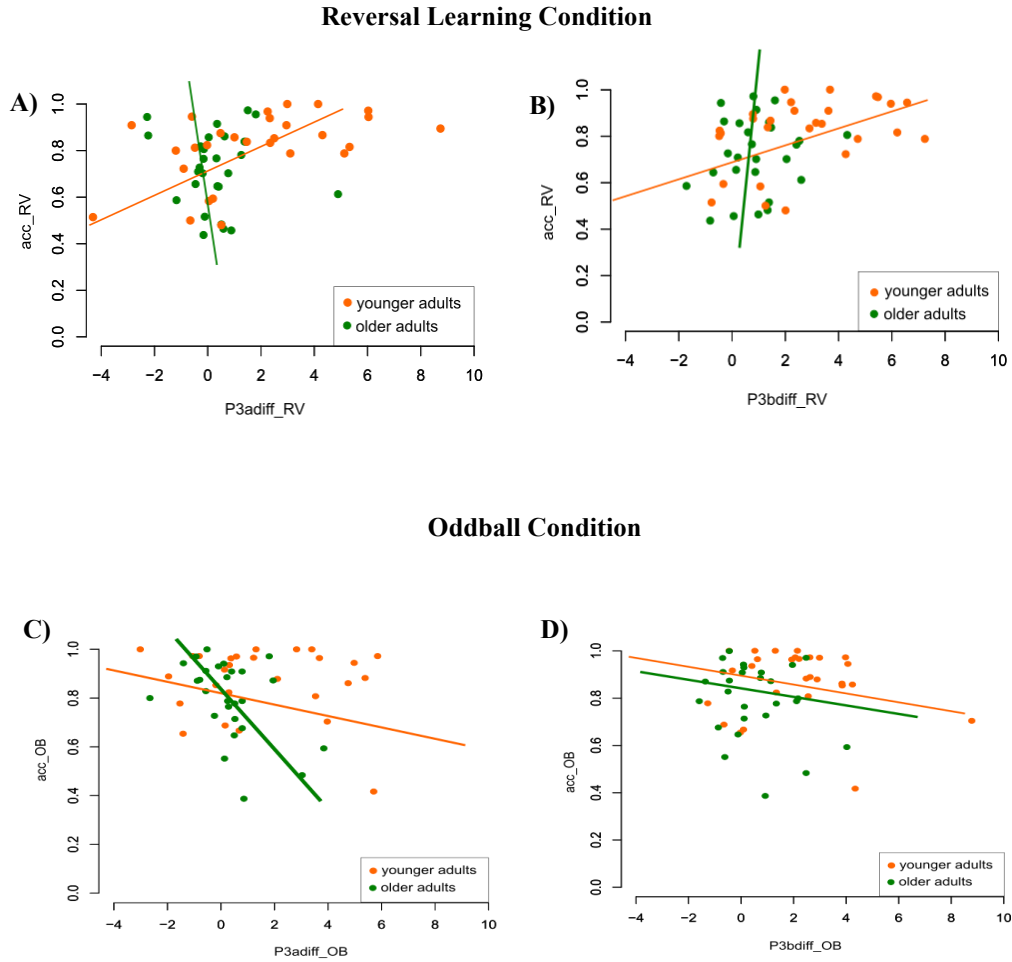
### *ERP-behavior correlations*

To investigate relationships between the behavioral and electrophysiological data across individuals, we ran Pearson correlation analyses. As shown in figure 6A, using difference values for the effect of trial type (rare-common) for the P3a component, we found a significant positive correlation between younger adults' accuracy on the rare+1 trial type and the P3a amplitude difference between common and rare trial types in the reversal learning condition ( $r=0.497$ ,  $p=0.028$ ). In contrast, we found no such correlation between behavior and the P3a amplitude in the reversal condition for the older adults ( $r=-0.093$ ,  $p=1.000$ ). For the oddball condition, we found no correlation between P3a trial differences and behavior for younger adults ( $r=-0.117$ ,  $p=1.000$ ), but found a trending negative correlation for the older adults ( $r=-0.450$ ,  $p=0.076$ ). Follow up analyses revealed that younger adults' correlation between the P3a and behavior significantly differed from the same correlation in older adults ( $z = 2.174$ ,  $p=0.029$ ). In contrast,

younger and older adults' correlations between the P3a and behavior did not significantly differ for the oddball condition ( $z= 1.330, p=0.184$ ) (see Fig 6C).

For the P3b component we calculated the same trial type difference values (rare-common). Examining the correlations between these P3b differences values and behavior on the rare+1 trial type revealed a significant positive correlation for younger adults in the reversal learning condition ( $r=0.412, p=0.116$ ). In comparison, we found no significant correlation between the P3b and behavior in the reversal condition for older adults ( $r=0.182, p=1.00$ ) (see Fig 6B). Follow up analyses revealed that correlations did not differ significantly between age groups ( $z = 0.719, p=0.472$ ). Finally, we found no significant correlations between P3b differences values and behavior in the oddball condition (younger adults:  $r=-0.216, p=1.000$ ; older adults:  $r=-0.229, p=1.000$ ) (see Fig 6D).

Together, our correlation analyses reveal that younger adults show a positive relationship between trial difference values for the P3a and performance on the rare+1 trial type in the reversal learning condition only. In contrast, older adults show no relationship between the P3a or P3b and subsequent behavior in the reversal learning condition. Nevertheless, older adults showed a trending negative relationship between their P3a and performance in the oddball condition, revealing that older adults with a smaller difference across trial types in their P3a amplitude showed better performance in the oddball condition. We found no other significant relationships between either component and performance on rare+1 in the oddball condition.



**Figure 5.** Scatterplots of correlations between average accuracy on the rare+1 trial type and average P3a or P3b trial difference (rare-common) at the participant level for both the reversal learning condition (top) and oddball condition (bottom). **A)** Correlation between average accuracy on the rare+1 trial type and P3a component trial difference value in the reversal learning condition. **B)** Correlation between average accuracy on the rare+1 trial type and P3b component trial difference value in the reversal learning condition. **Bottom:** Scatterplot of correlations for the oddball condition. **C)** Correlation between the accuracy on the rare+1 trial type and P3a component trial difference value (rare-common) in the oddball condition. **D)** Correlation between the accuracy on the rare+1 trial type and P3b component trial difference value (rare-common) in the oddball condition. For all plots, younger adult data is in orange, and older adults in green.

## Discussion

During healthy aging, the ability to make goal-directed decisions that rely on an internal model of the environment, has been shown to decline (Bolenz et al., 2019; Eppinger et al., 2013; Ruel et al., 2022). Recent theoretical work suggests that these deficits arise from older adults' underlying difficulty to learn and represent the state spaces of cognitive tasks (Eppinger et al., 2023). From a computational (reinforcement learning, RL) perspective, older adults' state space representation deficits could be a result of two different processes: 1) an inaccurate representation of states, actions, or transition probabilities or 2) an inefficient updating of these transitions. While both explanations are consistent with what has been shown behaviorally (Bolenz, et al., 2019; Eppinger et al., 2015; Hämmerer et al., 2019; Ruel et al., 2022), in the current study, we designed a novel paradigm to specifically test the latter explanation; inefficient updating.

In a novel paradigm, younger and older adults completed a decision-making task under two conditions: a reversal learning condition and an oddball condition. While the conditions differed in terms of their updating requirements following surprising outcomes, they were made identical with respect to their perceptual input. Therefore, participants had to recognise surprising outcomes in both conditions, but only update their internal representation following surprising outcomes in the reversal learning condition. Note that we refer to oddball trials in the oddball condition and to the reversal trials in the reversal learning condition as belonging to the rare trial type. The rare+1 trial type was consequently defined as trials immediately following the rare trials in each condition. Finally, the common trial type included all other trials.

Behavioral results revealed that both age groups showed reduced accuracy in the reversal learning condition as compared to the oddball condition. However, in line with past results (Hämmerer et al., 2019; Weiler et al., 2008), we found that older adults showed greater deficits in the reversal learning condition in comparison to younger adults. Specifically, older adults demonstrated marked slowing and error rates on common, versus rare trial types. Yet, they remained, overall, slower than young adults on the rare trial type. Taken together, in line with past findings by Hämmerer et al., (2019), these findings may reflect that older adults were overestimating the probability of reversals. Alternatively, this finding may reflect age differences



in choice randomness. Both these explanations may be manifesting in generally more stochastic responding, which, in the context of the current task, may explain older adults' apparent boosts in rare trial type performance. Future computational work should be used to disambiguate between these two possibilities. As predicted, age differences were less pronounced in the oddball condition (see figure 2). That is, although older adults' reaction times were significantly slower than that of younger adults, accuracy in the oddball condition did not differ between age groups.

Together, our behavioral results suggest that older adults showed greater impairments in the reversal learning condition compared to the oddball condition than younger adults did. Yet, reversals do not seem to have significantly affected the performance of either age group.

To investigate the neural dynamics underlying age-related changes in the ability to update the internal model of the environment following surprising outcomes, we examined EEG responses to outcome stimuli on rare trial types in comparison to common trial types in both conditions. Specifically, we focused on two ERP components that have been shown to reflect updating of internal models in response to surprising outcomes during cognitive tasks: the P3a, P3b.

In line with our predictions, we found a larger P3a component in the reversal learning condition as compared to the oddball condition as well as a larger P3a response to rare outcomes in comparison to common outcomes for younger adults. In contrast, older adults showed no significant difference in their P3a response across common and rare trial types but maintained a larger P3a component in the reversal learning condition. These findings indicate that younger adults were able to differentiate between common and rare trial types in both conditions (Polich 2007; Menon et al., 1997). In contrast, older adults showed no differentiation between common and rare trial types, suggesting that for the older age group, rare trials did not elicit a larger neural P3a response than common trials did. Consequently, we propose that older adults may not have experienced the rare trial type as surprising.

In contrast to our prediction that the P3b would be larger in the reversal learning condition compared to the oddball condition, reflecting internal model updating in this condition (Donchin, 1981; Donchin & Coles, 1988; Polich, 2007), we found no significant difference in the P3b component between conditions. Yet, both age groups showed a larger P3b response to rare

as opposed to common outcomes. Further, a significant interaction between trial type and age group revealed that younger adults showed a larger difference in P3b response between common and rare trial types, whereas older adults exhibited more similar (albeit, nevertheless significantly different) responses across trial type. This finding indicates that both age groups had a differentiated representation of common and rare trial types, with younger adults showing a more differentiated representation than older adults.

To better understand the significance of our ERP findings and their relationship to behavior on subsequent trials in each condition, we ran correlations between trial difference waves (rare-common) and accuracy on the rare+1 trial type. We found that the difference waves for the P3a related to younger adults' performance on subsequent (rare+1) trials, but only in the reversal learning condition. Specifically, younger adults who showed a larger differentiation between common and rare trials in the P3a component in the reversal learning condition also showed better performance on rare+1 trials in this condition. Therefore, performance following reversals may be related to younger adults' ability to discriminate between common and rare outcomes.

Interestingly, this finding is in line with results from a study by Nassar et al., (2019). In their study, they find no significant difference in the P300 component elicited in an oddball condition and changepoint condition (in which the structure of the task changes following surprising outcomes but creates a new structure every time). Nevertheless, Nassar et al., (2019) find a relationship between the amplitude of the P3 and learning. In their changepoint condition, larger P300 responses predicted increased learning. In contrast, in the oddball condition, larger P300 responses predicted reduced learning. In contrast to the younger adults, we found no significant relationship between P3a or P3b difference values and behavior in the reversal learning condition for older adults. This pattern of results suggests that older adults' behavioral difficulties in the reversal learning condition may arise from a deficit in discriminating between common and rare outcomes.

To summarize, we provide preliminary evidence that older adults have deficits in updating their internal representation during goal-directed reversal learning. However, in contrast to our prediction that internal model updating should be accompanied by a larger P3b, we found

no consistent increase in the P3b in a reversal learning condition. Instead, we found that the ability to differentiate common and rare outcomes during reversal learning is related to better performance on subsequent trials in younger but not older adults. Therefore, we suggest that the P3a may be involved in updating but does so by mediating the subjective response to surprise, which in turn affects behavior. We therefore suggest that older adults' deficits in discriminating between surprising outcomes that can be safely ignored, from those that must be used to update their internal model may explain their behavioral deficits during reversal learning.

One alternate explanation however remains. That is that older adults could be demonstrating a shift in their strategy, towards one in which they expect reversals or instability in the environment more, as a baseline, relative to younger adults. Older adults could therefore have had an accurate representation of the task, but simply were expecting reversals to occur more frequently, leading to the current pattern of behavior. This possibility should be examined in future research.

Although confirming this finding requires future studies to examine these effects using additional computational and neurological methods, we nevertheless suggest that in line with the diminished state space theory of cognitive aging (Eppinger et al., 2023), age-related deficits in the updating of state spaces and provide a starting point for future research. Future studies could adopt a single-trial regression approach to better examine the relationship between the P3b and behavioral adjustments or opt for an analysis technique that is not restricted in time, such as time-frequency analyses (Blanco et al., 1995).

## References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bolenz, F., Kool, W., Reiter, A. M., & Eppinger, B. (2019). Metacontrol of decision-making strategies in human aging. *Elife*, *8*, e49154.
- Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., De Gardelle, V., Castañón, S. H., & Summerfield, C. (2014). Adaptive gain control during human perceptual choice. *Neuron*, *81*(6), 1429-1441.
- d'Acremont, M., & Bossaerts, P. (2016). Neural mechanisms behind identification of leptokurtic noise and adaptive behavioral response. *Cerebral Cortex*, *26*(4), 1818-1830.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.
- Donchin, E. (1981). Surprise!... surprise?. *Psychophysiology*, *18*(5), 493-513.
- 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*, 253.
- Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243-1255.
- 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.
- Jepma, M., Brown, S. B., Murphy, P. R., Koelewijn, S. C., de Vries, B., van den Maagdenberg, A. M., & Nieuwenhuis, S. (2018). Noradrenergic and cholinergic modulation of belief updating. *Journal of Cognitive Neuroscience*, *30*(12), 1803-1820.
- Jodo, E., & Kayama, Y. (1992). Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalography and clinical neurophysiology*, *82*(6), 477-482.

- Kopp, B., Mattler, U., Goertz, R., & Rist, F. (1996). N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming. *Electroencephalography and clinical Neurophysiology*, 99(1), 19-27.
- Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. *elife*, 8, e46975.
- Nassar, M. R., Bruckner, R., Gold, J. I., Li, S. C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications*, 7(1), 11609.
- Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, 30(37), 12366-12378.
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, 51, 195-203.
- Polich, J., & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology*, 60(2), 172-185.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, 118(10), 2128-2148.
- Pontifex, M. B., Hillman, C. H., & Polich, J. (2009). Age, physical fitness, and attention: P3a and P3b. *Psychophysiology*, 46(2), 379-387.
- Ruel, A., Bolenz, F., Li, S. C., Fischer, A., & Eppinger, B. (2022). Neural evidence for age-related deficits in the representation of state spaces. *Cerebral Cortex*, 33(5), 1768-1781.
- Samanez-Larkin, G. R., Worthy, D. A., Mata, R., McClure, S. M., & Knutson, B. (2014). Adult age differences in frontostriatal representation of prediction error but not reward outcome. *Cognitive, Affective, & Behavioral Neuroscience*, 14, 672-682.
- Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions?. *Trends in cognitive sciences*, 19(1), 27-34.
- Weiler, J. A., Bellebaum, C., & Daum, I. (2008). Aging affects acquisition and reversal of reward-based associative learning. *Learning & memory*, 15(4), 190-197.

## CHAPTER 5: GENERAL DISCUSSION

The overarching goal of this dissertation was to (1) explore the mechanisms underlying age-related changes in decision-making and (2) to develop a theoretical framework that characterizes these changes. In the following sections, I outline the main findings from the two empirical studies (chapters 2 & 4) and summarize the theoretical paper (chapter 3). Additionally, I discuss the main contributions of each chapter. Finally, I discuss the implications and limitations of this work and suggest potential future directions.

### Summary and contributions of main findings

#### Chapter 2

**Summary.** In chapter 2 (empirical study 1), I examined if older adults would demonstrate a greater contribution of model-based decision-making when the demands on representing the task environment were reduced. To do so, I developed a modified version of a Markov two-step task in which I manipulated the demands on representing the transition probabilities linking the first and second stages of the task. By using a combined electrophysiological (EEG) and computational (RL) approach, I found that younger, but not older adults benefited from more differentiated transition probabilities. Specifically, results of a modified hybrid computational RL model revealed that in contrast to younger adults, older adults were less able to accurately represent the transition probabilities, and therefore failed to predict the value of the upcoming choice options. Neural (EEG) analyses corroborated past work showing that the P300 component covaried with the degree to which participants update their representation in the probabilistic transition structure (Eppinger et al., 2017). Importantly, using a single trial EEG regression approach (see Fischer & Ullsperger, 2013), I found that younger adults' P300 component related to their ability to represent the relationship between their choice and the value of the resulting choice options. In contrast, older adults' data showed no such relationship. This pattern of results suggests that older adults may experience difficulties in predicting the value of upcoming choice options, which may be a downstream consequence of their diminished representations of the task structure. Interestingly, the same single-trial EEG

regression analysis examining the FRN/RewP and trial-by-trial rewards prediction errors (RPEs) demonstrated that both age groups showed intact reward processing. That is, both younger and older adults had a distinct representation of rewarded and unrewarded outcomes which, in turn was related to their trial-by-trial RPEs.

Taken together, older adults appear to have specific deficits in representing the probabilistic transition structure of decision-making tasks and in predicting the value of upcoming choice options. Therefore, in line with previous work, I suggest that age-related changes in decision-making may be due to impairments in the representation of task environments.

**Contributions.** One of the major contributions of the first empirical paper (see chapter 2), is its use of a single trial EEG regression approach (Fischer & Ullsperger, 2013). By applying a single trial EEG regression analysis (based on Fischer & Ullsperger, 2013), I was able to examine latent processes that I would not have been able to directly examine otherwise. Specifically, these multiple robust regression analyses allowed me to examine trial-by-trial correlations between ERP components (P300 and FRN/RewP) and time-varying parameters extracted from a computational RL model (i.e., prediction errors). This is the first study that directly examined previous predictions that the P300 covaries with value prediction errors (VPEs) while the FRN/RewP covaries with reward prediction errors (RPEs). Consequently, this analysis allowed me to provide novel evidence that older adults experience specific deficits with transition structures during decision-making, while their ability to adjust their behavior as a function of reward feedback remains intact.

Based on the importance and novelty of the findings described above, a related contribution of this paper is the role it played in developing the diminished state space theory (theoretical paper presented in chapter 3). Alongside several other empirical findings, results from this study helped provide a strong argument that age-related impairments in model-based decision-making result from deficits in the representation of state spaces.

### Chapter 3

**Summary.** In chapter 3 (theoretical paper), I propose the diminished state space theory of human aging which relies on findings from the first empirical paper, as well as several other critical findings on the age-related changes in decision-making.

The main argument in this theoretical paper is that older adults' difficulties engaging in goal-directed decision-making can be explained by an underlying deficit in the representation of state spaces of cognitive tasks. A state space, in this theory, is defined as the abstract representation of the structure of the task which includes 1) the decision options (stimuli) that are available in a given cognitive task, 2) the contingencies between the options and the corresponding actions and 3) the outcomes that follow these actions (see Wilson et al., 2014). In the paper, we present empirical, computational, and neurobiological evidence to support this argument.

In the literature review section of the paper, we provide evidence that older adults seem to have hard limits in their ability to engage in model-based learning which relies on the creation and maintenance of an abstract representation the task structure. For example, even when additional incentives are provided (Bolenz et al., 2019) or when representational demands are reduced (Ruel et al., 2022; see chapter 2), older adults still show difficulties engaging in model-based decision-making. From a computational standpoint, we propose two non-mutually exclusive mechanisms to explain older adults' deficits in state space representations. First, older adults may have simplified representations of the state space. By simplifying the rules that link one state to the next, or by pruning part of the state space, older adults may end up with an inaccurate representation. Second, they may struggle to update the transition probabilities based on feedback obtained during decision-making. Finally, from a neurobiological perspective, deficits in the representation and updating of state spaces seem to be (at least in part) explained by the functional decline of the hippocampus and orbitofrontal cortex (OFC), respectively.

Together, the proposed diminished state space (DSS) theory is shown to explain age-related changes in decision-making and more broadly, may extend to cognitive control.

**Contributions.** The main contribution of the diminished state space (DSS) theory is its success in synthesizing a large body of work that had yet to be expressed in a single theory. Although several research groups around the world have directly or indirectly suggested that



older adults may have deficits representing abstract task structures (see Chan et al., 2021; Hämmerer et al., 2019; Nassar et al., 2016; Rmus et al., 2019), this theory provides the field with a concise description of what we know about older adults' deficits and sets the tone for future work in cognitive aging. Future work should examine our proposed computational mechanisms and neurobiological predictions. Specific implementations are proposed in the future directions section below.

## Chapter 4

**Summary.** One of the major benefits of model-based decision-making is its flexibility in changing environments. Yet, these changes need to be incorporated in the internal representation of the task. Therefore, in chapter 4 (study 2), I examined one of two computational explanations proposed in the diminished state space theory regarding older adults' representational deficits: difficulties *updating* their internal representation.

By designing a simple paradigm which allowed me to dissociate updating from surprise, I examined if older adults differ from younger adults in their ability to discriminate between outcomes that must be used to update their internal model (i.e., reversals) from those that can be safely ignored (i.e., oddballs). In one condition (reversal learning condition), a surprising outcome indicated that the transition structure of the task has changed, and therefore dictated the need to update the internal model. In contrast, in the oddball condition, surprising outcomes, which were perceptually identical to those in the reversal condition, could be safely ignored. Our behavioral results revealed that older adults show greater behavioral deficits in the reversal learning condition as compared to younger adults. However, in contrast to our predictions, we found no difference in the P3b across conditions for either age group. Nevertheless, we did find that the P3a systematically related to better performance on trials immediately after reversals in the reversal learning condition for younger, but not older adults. We interpret these findings in line with previous work by Nassar et al. (2019) which demonstrates that the P300 reflects the need to increase or decrease learning based on the statistical context. That is, we suggest that older adults' behavioral deficits during reversal learning may be explained by their difficulties in discriminating between surprising outcomes that necessitate an update to the internal model, from those that do not.

Overall, we provide preliminary evidence that older adults experience difficulty in correctly identifying and learning from surprising outcomes in changing environments. In line with the diminished state space theory of human aging, we therefore propose that older adults have difficulty updating their internal model during decision-making.

**Contributions.** One main contribution of this theoretical paper is its experimental design which was simplified in order to focus on updating processes. That is, in line with the DSS theory suggesting that additional states, actions and transitions between states may exacerbate older adults' difficulty to represent the whole state space, I reduced the complexity of the task to include only two deterministic transitions between two states. This simple design therefore allowed me to test for age-related differences in updating in a task that had been stripped of additional factors which could make updating more difficult (i.e., additional states and/or probabilistic transitions).

Another benefit of designing a simple paradigm is the ease with which it can be modified in order to answer additional research questions. For instance, by making the state space more complex (i.e., by adding more states, actions, or transitions), the same design could be used to examine if manipulating the size of state space significantly affects older adults' ability to update their internal model. In line with the diminished state space theory, I would predict that older adults' deficits in updating would be further enhanced in a large space as compared to a small space condition. Alternatively, the volatility of the environment could be increased, by decreasing the number of trials between reversals. Similar to increasing the size of the state space, I would predict that increasing the volatility would lead to more pronounced deficits in older adults. Finally, the task could be used to examine updating processes in children who have been suggested to experience similar difficulties with updating (Decker et al., 2016; Potter et al., 2017; Smid et al., 2020).

## **Implications**

A major implication of the research described in this dissertation is its relevance in real world situations. As western societies become increasingly reliant on technology, understanding why current designs remain barriers for older adults is of paramount importance. Decades of research

have and continue to determine how we can modify current technology to help older adults who experience age-related sensory and cognitive decline (Di Giacomo et al., 2019; Gitlow, 2014; Tyler, De George-Walker & Simic, 2020). However, there is an additional underlying deficit that must also be considered: older adults' difficulties in learning and making decisions within the hierarchical information structures (i.e., menus) inherent to most modern technologies (Ziefle & Bay, 2004; 2006). The current dissertation takes a step towards understanding these deficits by demonstrating that older adults experience marked difficulties representing the state space of cognitive tasks which can be thought of as conceptually comparable to the hierarchical menu structure used in many technologies.

In addition to identifying that older adults' representational deficits may explain difficulties engaging with technology in the real world, we propose two solutions that may help circumvent these difficulties.

First, possibly the simplest solution, is to provide technology users a map of the state space (i.e., a map of the menu structure). Past findings have demonstrated that external prompts for action or environmental support help older adults' performance (Mayr, Spieler & Hutcheon, 2015). A map could significantly help older adults' decision-making by allowing them to offload some of the computational costs associated with representing the state transitions to an external aid. Although there is significant evidence showing that when external support is removed, older adults show noticeable performance deficits, there is also some evidence that external support may even benefit older adults once the source of information is removed (Lindenberger & Mayr, 2014). For instance, in the last block of a paradigm in which environmental sources of information are initially present and removed over time, older adults' accuracies remain very high (Lindenberger & Mayr, 2014). Although future research will need to investigate the features that are most helpful to older adults, there is preliminary evidence that including the agent's current location in a map may also be beneficial. According to previous research from the field of interface design, maps that reflect the user's current location relative to other submenus may significantly improve older adults' ease in using complex devices (Zeiffler & Bay, 2006). Lindenberger and Mayr (2014), provide additional recommendations, suggesting that external cues should be compatible (i.e., in line with the user's goal) and distinctive (i.e., each cue should suggest only one action).

Second, simplifying the state space by reducing the number of available states and actions may significantly help older adults who have difficulties representing them. Reducing the complexity of the state space by reducing the planning depth, may help older adults engage in more complex decision-making strategies. While this solution remains more speculative, in the DSS paper (see chapter 3), we allude to a computational explanation which suggests that older adults may naturally try to simplify complex state spaces. Therefore, by starting out with simpler state spaces, we may see that older adults experience less difficulty in accurately representing all states, actions and transitions that comprise them.

A second implication of the current dissertation pertains to its applicability to clinical research. While some work has begun to investigate the impact of clinical aging on decision-making, much work remains to develop a clear understanding of the cognitive deficits that accompany conditions such as mild cognitive impairment (MCI) and Alzheimer's disease. For instance, past work has shown that, similar to healthy aging adults, Alzheimer patients experience difficulties with reversal learning (Freedman & Oscar-Berman, 1989). Further, Parkinson's patients have been shown to have deficits in goal-directed decision-making, while their habitual decision-making remains intact. Therefore, by providing a clearer description of the impact of healthy aging on decision-making, we can begin to better quantify the effects of clinical aging on decision-making. Further, by understanding the degree to which specific decision-making abilities are affected in healthy aging, and the mechanism by which these deficits arise, we may be able to help identify the progression of different clinical disorders. This, in turn, could help earlier detection and potential treatment. For example, recent work by Ahmed et al., (2022), has demonstrated that we may be able to disentangle frontotemporal dementia (FTD) from Alzheimer's (AZ) disease using a reversal paradigm. They found that the AZ group demonstrated significantly greater deficits in stimulus-association learning than the FTD group, therefore suggesting that such paradigms may help discriminate between two similar diagnoses and help inform treatment plans. Similarly, in a review paper, Santos de Siquiera et al., (2017) demonstrate that both probable Alzheimer's patients and individuals with MCI experienced impairments in decision-making in ambiguous and risky contexts. However, Alzheimer's patients showed stronger impairments.

Thus, while the work in the current dissertation does not directly examine clinical populations, these findings may nevertheless help advance clinical research.

## **Limitations**

One main limitation of the present dissertation is the small sample sizes used in both empirical studies. That is, the sample sizes used do not allow for the assessment of factors that may lead to inter-individual differences in decision-making.

For instance, several psychiatric, psychological, or neurological conditions such as schizophrenia (Culbreth et al., 2016), depression (Heller et al., 2018), anxiety (Bishop & Gagne, 2018) and gambling disorder (Wyckmans et al., 2019), as well as aging disorders such as mild cognitive impairment (MCI; Zamarian, Weiss & Delazer, 2011), have been shown to impact decision-making processes including model-based learning. In comparison to healthy older adults, individuals with MCI have been shown to experience greater difficulties in learning from feedback and adapting to changes in the decision-making environment (Zamarian, Weiss & Delazer, 2011), abilities that were required and closely examined in the empirical studies reported in this dissertation. Therefore, if any of the older adults included in the final samples were diagnosed or had MCI symptoms, it may have led to larger differences between them and the younger adults; leading to the false conclusion that healthy older adults have larger deficits than they actually do.

Although I took precautions to ensure that all participants included in both empirical studies were free of any psychological or neurological condition that could impact their performance, participants were not clinically evaluated for these conditions. Therefore, since many individuals remain undiagnosed or fail to report their diagnosis (see Pelletier et al., 2017), there is a possibility that a few participants may have diagnoses that could have affected their performance. There is also evidence that sub-clinical traits can result in changes in decision-making. For instance, impulsivity (Raio, Konova & Otto, 2020), and even transient changes in factors such as stress (see Otto et al., 2013; Schwabe & Wolf, 2011), can shift decision control away from model-based learning. As the studies in this dissertation did not account for individual differences in stress and impulsivity, these factors may also be playing a small role in the reported findings.

Ultimately, the studies included in this dissertation do not assess the potential clinical and sub-clinical conditions and traits that have previously been shown to impact the engagement in model-based decision-making. Rather, they focus on examining and explaining the general changes that occur with (reported) healthy aging. Nevertheless, considering these individual differences, the current findings should be replicated as we continue to investigate age-related changes in decision-making.

As such, I am currently working in collaboration with Dr. Moritz Koester from Universität Regensburg, and Mirjam Marx from the Freie Universität Berlin to run a replication of the empirical study described in chapter 4. By doing so, I aim to examine the validity of the findings, as well as potentially beginning to examine some of the traits that may lead to individual differences in younger as well as older adults' abilities.

A second limitation of the present dissertation pertains to the methodological approach. In both empirical studies, I used event-related potentials (ERPs) as a primary method to investigate the neural mechanisms behind age-related changes in decision-making. However, there is one considerable limitation of ERP analyses. As suggested by its name, ERP analyses are restricted to examining event-related activity, and therefore do not leverage all the information contained within the EEG signal. Specifically, ERP analyses ignore all non-phase-locked signals as they focus on neural activity in short time windows (in the order of milliseconds), following an event which can be the presentation of a stimulus or a response. However, there is much more to the EEG signal, and these patterns of neural activity may be of particular importance in decision-making. For instance, although several empirical studies (including those in this dissertation) have shown that the process of updating the internal representation of a task typically begins when a stimulus or outcome indicates that the environment has changed (Eppinger et al., 2017; Eppinger et al., 2013; Gläscher et al., 2010), the updating process may span a few seconds, making the ERP technique suboptimal to capture the process in its entirety. Additionally, a lack of significant condition effects in ERP analyses do not necessarily mean that there are no differences in the reported comparison. For example, non-phased locked dynamics that are task modulated are often not observable in the ERP, as the averaging process inherent to ERP analyses most often averages out this variance. In the Future Directions section below, I

discuss how wavelet analyses can help circumvent these issues and be used to compliment standard ERP analyses.

Finally, although this dissertation focuses on age-related changes in the ability to create and update an abstract representation of a cognitive task, neither empirical study directly examined neural representations. Instead, I assume that, based on previous work (Bolenz et al., 2019; Eppinger et al., 2013; Eppinger et al., 2013) the experimental manipulations in the first study (see chapter 2) impacted the demands on representing the internal model of the task. Similarly, in the second empirical study (see chapter 4), I assume that the developed task has simplified the initial demands on representing the task environment, resulting in a paradigm that is aimed at specifically examining updating. Therefore, future work should attempt to directly examine neural representations. One method which shows great promise in doing so is representational similarity analyses (RSA); a multivariate technique that allows the examination of similarities among neural activity patterns (in fMRI or EEG data) generated by different conditions (e.g., different demands on representational abilities). Details of this method are discussed in the Future Directions section.

## **Future Directions**

While the current dissertation provides noteworthy findings that have advanced our understanding of the specific age-related deficits in goal-directed decision-making, there remain several open questions. In what follows, I outline future directions both in terms of methodological approaches and research questions.

First, future research should continue to adopt different analysis techniques to complement traditional ERP and computational RL analyses. One such approach is single trial regression analyses, used in the first empirical study, which allows for the examination of latent processes such as the relationship between time varying computational parameters and neural activity. Another technique is time-frequency (wavelet) analysis. The unique advantage of time-frequency approaches is that neural activity is analysed in the frequency domain as opposed to

the time domain. Further, this approach allows for an analysis over longer periods of time when compared to ERPs (see Kolev et al., 2005).

Within the decision-making literature, previous research has successfully demonstrated that there are at least two frequency bands that contribute to the P300: the delta and theta band (Başar-Eroglu et al., 1992; Harper, Malone & Iacono, 2017), which have been shown to relate to creating and updating internal representations, respectively. For instance, the theta band has been shown to covary with reward prediction errors across several studies (Cavanagh et al., 2012; Cavanagh & Frank, 2014), while the delta band has been shown to reflect response-related demands (Harper, Malone & Iacono, 2017). Therefore, using time-frequency analyses may be another (and potentially more informative) way to examine mechanisms such as those that support updating internal models. Specifically, as updating processes can potentially last longer than the time windows examined using ERPs and/or not be bound to any specific stimulus or response, wavelet analyses may be particularly informative.

As part of my collaboration with Dr. Moritz Köster and Mirjam Marx, in a follow-up study to the experiment described in chapter 4, we will be applying time-frequency analyses to examine the neural oscillations elicited following surprising outcomes in the reversal learning condition.

A second method that shows great promise for the study of neural representations is inspired from multi-voxel pattern analyses (MVPA) used in MRI research (Schuck et al., 2016). Namely, representational similarity analysis (RSA) has been used to study differences in cognitive representations reflected in EEG data (Freund et al., 2021). This multivariate technique allows for the decomposition of control-related neural activity (EEG or fMRI data) into measures that better correspond to representational components of theories. Past work using this technique has been successful in examining similarity structures across neural activity patterns of different conditions (see He et al., 2021; Kriegeskorte et al., 2008). Therefore, RSA may be used as a method to examine the hypothesis that older adults experience representational deficits. That is, by applying RSA to experimental paradigms designed to examine the neural processes behind the creation and updating of state spaces, we may be able to further investigate age-related differences in these abilities.



A second future direction pertains to the theoretical manuscript in this dissertation (see chapter 3). Future research is needed to determine if the diminished state space (DSS) theory can be generalized to other cognitive domains such as cognitive control. As discussed in the DSS theory paper, several cognitive control tasks such as task switching, or the Stroop task, require the maintenance of several task sets (stimulus-action pairings) in order to achieve good performance (Musslick & Cohen, 2021). Therefore, the DSS theory would predict that older adults' performance on these tasks should be explained by their greater difficulties representing and maintaining an accurate representation of the task, with larger state spaces taxing their abilities to a greater degree. A handful of studies examining age-related differences in task switching performance suggest that this prediction may hold true (Verhaeghen, 2011; Wasylshyn et al., 2011). That is, the ability to represent and update two task sets declines with age, not the specific ability to switch between two tasks (Wasylshyn et al. 2011). Therefore, consistent with the DSS theory, it seems plausible that age-related difficulties in goal-directed behavior, even in the context of cognitive control, may arise due to deficits in the representation and/or updating of state spaces. Future research could examine these predictions using task switching paradigms by manipulating the demands on representing the state space (i.e., changing the number of task sets; or manipulating the transition probabilities within task sets).

As part of another ongoing collaboration with a Bachelor's student, Kimia Motevalli, I am currently examining if the DSS theory generalizes to older adults' performance in task switching. To do so, I have created a novel task switching paradigm in which participants must switch between three tasks. In line with the DSS theory, we predict that older adults will show greater difficulties switching between tasks in a condition that includes a larger number of task sets (global switch costs) as opposed to switching between tasks at the trial level (local switch costs).

A second future direction arising from the DSS theory is our suggestion that observed deficits in older adults in the representation and updating of state spaces are due to the functional decline of the orbitofrontal cortex (OFC) and hippocampus. By using fMRI with simultaneous EEG, we may gain unique insight on the spatial and temporal dynamics underlying changes in decision-making. Finally, as mentioned above, novel analysis techniques such as MVPA and RSA could be applied to examine the OFC and hippocampus and their respective roles in state space representations required for model-based decision-making.

A final future direction regarding the overall proposal of older adult's diminished state space representations is an examination of the relative contribution of other mechanisms which could result in these impoverished representations. Based on a large body of work demonstrating that older adults have deficits in learning (see Finucane et al., 2005; Peters et al., 2007; Rolison et al., 2012), one possibility is that a deficit in learning in complex goal-directed environments may explain (in part) the reason older adults experience diminished state space representations. Another possibility is that older adults experience deficits retrieving information that has been previously learned (see Bäckman et al., 1997; Gilsky et al., 2001), which in turn could contribute to an impoverished representation. Future research should therefore examine the relative contributions of these mechanisms in order to help provide a complete mechanistic explanation as to how diminished state spaces may arise in older adults.

## **Conclusion**

In conclusion, the present dissertation contributes to the field of decision-making and more broadly, cognitive aging, by providing important insights to our understanding of the neurocomputational mechanisms behind age-related changes in decision-making, and providing a novel theory of human cognitive aging. In chapter 2, I provide the first direct evidence for age-related deficits in the ability to represent probabilistic transitions and predict the value of upcoming choice options. In chapter 3, I propose a novel theoretical explanation (the diminished state space theory) to the deficits experienced by older adults which takes into consideration the empirical work described here, among other previous studies on age-related changes in decision-making. Finally, in chapter 4, I provide preliminary evidence that older adults experience difficulties updating their internal model by showing that they have deficits identifying surprising outcomes during decision-making. By utilising novel analysis approaches such as single-trial regression analyses, findings from the current dissertation significantly advance our understanding of the mechanisms behind the age-related changes in decision-making. As we begin to develop a better understanding of these neuro-computational mechanisms, we can move towards creating devices that will be designed with the cognitive abilities of individuals of all ages in mind. This will help us move towards a more inclusive tech-supported society.

## References

- Ahmed, K., Mitchell, D. G., Blair, M., Coleman, K., Pasternak, S. H., Ruiz-Garcia, R., & Finger, E. (2022). Disentangling Reversal-learning Impairments in Frontotemporal Dementia and Alzheimer Disease. *Cognitive and Behavioral Neurology*, *35*(2), 110-122.
- Başar-Eroglu, C., Başar, E., Demiralp, T., & Schürmann, M. (1992). P300-response: possible psychophysiological correlates in delta and theta frequency channels. A review. *International journal of psychophysiology*, *13*(2), 161-179.
- Bishop, S. J., & Gagne, C. (2018). Anxiety, depression, and decision making: a computational perspective. *Annual review of neuroscience*, *41*, 371-388.
- Bolenz, F., Kool, W., Reiter, A. M., & Eppinger, B. (2019). Metacontrol of decision-making strategies in human aging. *Elife*, *8*, e49154.
- Bornstein, A. M., & Daw, N. D. (2013). Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLoS computational biology*, *9*(12), e1003387.
- Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience & Biobehavioral Reviews*, *26*(7), 809-817.
- Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., ... & Reed, B. R. (2001). Context processing in older adults: evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental*
- Braver, T. S., Satpute, A. B., Rush, B. K., Racine, C. A., & Barch, D. M. (2005). Context processing and context maintenance in healthy aging and early stage dementia of the Alzheimer's type. *Psychology and aging*, *20*(1), 33.
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences*, *18*(8), 414-421.
- Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220-238.
- Chan, S. C., Schuck, N. W., Lopatina, N., Schoenbaum, G., & Niv, Y. (2021). Orbitofrontal cortex and learning predictions of state transitions. *Behavioral neuroscience*, *135*(4), 487.

- Craik, F. I., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In *Aging and cognitive processes* (pp. 191-211). Boston, MA: Springer US.
- Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M., & Barch, D. M. (2016). Reduced model-based decision-making in schizophrenia. *Journal of abnormal psychology, 125*(6), 777.
- Czaja, S., & Schulz, R. (2006). Innovations in technology and aging introduction. *Generations, 30*(2), 6-8.
- Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current opinion in neurobiology, 16*(2), 199-204.
- 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.
- Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology, 18*(2), 185-196.
- de Siqueira, A. S. S., Yokomizo, J. E., Jacob-Filho, W., Yassuda, M. S., & Aprahamian, I. (2017). Review of decision-making in game tasks in elderly participants with Alzheimer disease and mild cognitive impairment. *Dementia and geriatric cognitive disorders, 43*(1-2), 81-88.
- Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological science, 27*(6), 848-858.
- Deserno, L., Huys, Q. J., Boehme, R., Buchert, R., Heinze, H. J., Grace, A. A., ... & Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences, 112*(5), 1595-1600.
- Di Giacomo, D., Guerra, F., Perilli, E., & Ranieri, J. (2020). Technophobia as emerging risk factor in aging: Investigation on computer anxiety dimension. *Health Psychology Research, 8*(1).
- 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.

- Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating?. *Behavioral and brain sciences*, *11*(3), 357-374.
- Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. *Behavioural brain research*, *103*(2), 123-133.
- Eppinger, B., Heekeren, H. R., & Li, S. C. (2015). Age-related prefrontal impairments implicate deficient prediction of future reward in older adults. *Neurobiology of aging*, *36*(8), 2380-2390.
- Eppinger, B., Walter, M., & Li, S. C. (2017). Electrophysiological correlates reflect the integration of model-based and model-free decision information. *Cognitive, Affective, & Behavioral Neuroscience*, *17*, 406-421.
- 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*, 253.
- Evangelista, N. D., O'Shea, A., Kraft, J. N., Hausman, H. K., Boutzoukas, E. M., Nissim, N. R., ... & Woods, A. J. (2021). Independent contributions of dorsolateral prefrontal structure and function to working memory in healthy older adults. *Cerebral Cortex*, *31*(3), 1732-1743.
- Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243-1255.
- Freedman, M., & Oscar-Berman, M. (1989). Spatial and visual learning deficits in Alzheimer's and Parkinson's disease. *Brain and cognition*, *11*(1), 114-126.
- 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.
- Gitlow, L. (2014). Technology use by older adults and barriers to using technology. *Physical & Occupational Therapy in Geriatrics*, *32*(3), 271-280.
- 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.

- Gunning-Dixon, F. M., & Raz, N. (2003). Neuroanatomical correlates of selected executive functions in middle-aged and older adults: a prospective MRI study. *Neuropsychologia*, *41*(14), 1929-1941.
- Hämmerer, D., Müller, V., & Li, S. C. (2014). Performance monitoring across the lifespan: still maturing post-conflict regulation in children and declining task-set monitoring in older adults. *Neuroscience & Biobehavioral Reviews*, *46*, 105-123.
- 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.
- Harper, J., Malone, S. M., & Iacono, W. G. (2017). Theta-and delta-band EEG network dynamics during a novelty oddball task. *Psychophysiology*, *54*(11), 1590-1605.
- Hascher, L., & Zacks, R. (1988). Working memory, comprehension and aging: a review and a new review. *The psychology of learning and motivation*, *22*, 193-225.
- He, J., Meyer, A. S., & Brehm, L. (2021). Concurrent listening affects speech planning and fluency: the roles of representational similarity and capacity limitation. *Language, Cognition and Neuroscience*, *36*(10), 1258-1280.
- Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature reviews neuroscience*, *5*(2), 87-96.
- Heller, A. S., Ezie, C. C., Otto, A. R., & Timpano, K. R. (2018). Model-based learning and individual differences in depression: The moderating role of stress. *Behaviour research and therapy*, *111*, 19-26.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, *109*(4), 679.
- Hunsaker, A., Nguyen, M. H., Fuchs, J., Djukaric, T., Hugentobler, L., & Hargittai, E. (2019). "He explained it to me and I also did it myself": How older adults get support with their technology uses. *Socius*, *5*, 2378023119887866.
- Jepma, M., Brown, S. B., Murphy, P. R., Koelewijn, S. C., de Vries, B., van den Maagdenberg, A. M., & Nieuwenhuis, S. (2018). Noradrenergic and cholinergic modulation of belief updating. *Journal of Cognitive Neuroscience*, *30*(12), 1803-1820.

- Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic regulation of learning rate in a dynamic environment. *PLoS computational biology*, *12*(10), e1005171.
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of experimental psychology: General*, *132*(1), 47.
- Kaplan, R., King, J., Koster, R., Penny, W. D., Burgess, N., & Friston, K. J. (2017). The neural representation of prospective choice during spatial planning and decisions. *PLoS biology*, *15*(1), e1002588.
- Kirasic, K. C., Allen, G. L., Dobson, S. H., & Binder, K. S. (1996). Aging, cognitive resources, and declarative learning. *Psychology and aging*, *11*(4), 658.
- Kolev, V., Falkenstein, M., & Yordanova, J. (2005). Aging and error processing: Time-frequency analysis of error-related potentials. *Journal of Psychophysiology*, *19*(4), 289-297.
- Kool, W., Cushman, F. A., & Gershman, S. J. (2018). Competition and cooperation between multiple reinforcement learning systems. *Goal-directed decision making*, 153-178.
- Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological science*, *28*(9), 1321-1333.
- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, *4*.
- Lengyel, M., & Dayan, P. (2007). Hippocampal contributions to control: the third way. *Advances in neural information processing systems*, *20*.
- Lindenberger, U., & Mayr, U. (2014). Cognitive aging: is there a dark side to environmental support?. *Trends in cognitive sciences*, *18*(1), 7-15.
- Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a "new view".
- 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.

- Metzler-Baddeley, C., Hunt, S., Jones, D. K., Leemans, A., Aggleton, J. P., & O'Sullivan, M. J. (2012). Temporal association tracts and the breakdown of episodic memory in mild cognitive impairment. *Neurology*, *79*(23), 2233-2240.
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature neuroscience*, *9*(8), 1057-1063.
- Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control. *Trends in Cognitive Sciences*, *25*(9), 757-775.
- Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. *elife*, *8*, e46975.
- Nassar, M. R., Bruckner, R., Gold, J. I., Li, S. C., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature communications*, *7*(1), 11609.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological bulletin*, *131*(4), 510.
- Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G., Holroyd, C. B., Kok, A., & Van der Molen, M. W. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 19-36.
- O'doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology*, *14*(6), 769-776.
- Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, *110*(52),
- Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *Journal of cognitive neuroscience*, *27*(2), 319-333.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and aging*, *17*(2), 299.



- Pelletier, L., O'Donnell, S., Dykxhoorn, J., McRae, L., & Patten, S. B. (2017). Under-diagnosis of mood disorders in Canada. *Epidemiology and psychiatric sciences*, 26(4), 414-423.
- Potter, T. C., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Developmental Cognitive Neuroscience*, 25, 272-280.
- Raio, C. M., Konova, A. B., & Otto, A. R. (2020). Trait impulsivity and acute stress interact to influence choice and decision speed during multi-stage decision-making. *Scientific reports*, 10(1), 7754.
- Rajah, M. N., & D'Esposito, M. (2005). Region-specific changes in prefrontal function with age: a review of PET and fMRI studies on working and episodic memory. *Brain*, 128(9), 1964-1983.
- 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.
- Resnick, S. M., Lamar, M., & Driscoll, I. (2007). Vulnerability of the orbitofrontal cortex to age-associated structural and functional brain changes. *Annals of the New York Academy of Sciences*, 1121(1), 562-575.
- Rmus, M., Ritz, H., Hunter, L. E., Bornstein, A. M., & Shenhav, A. (2019). Individual differences in model-based planning are linked to the ability to infer latent structure. *BioRxiv*, 723072.
- Ruel, A., Bolenz, F., Li, S. C., Fischer, A., & Eppinger, B. (2022). Neural evidence for age-related deficits in the representation of state spaces. *Cerebral Cortex*, 33(5), 1768-1781.
- Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, 2(3), 179-183.
- Salthouse, T. A. (1995). Influence of processing speed on adult age differences in learning. *Age*, 10(59), 55.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological review*, 103(3), 403.

- Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, *27*(47), 12860-12867.
- Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447), eaaw5181.
- Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, *91*(6), 1402-1412.
- Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. *Behavioural brain research*, *219*(2), 321-328.
- Simon, D. A., & Daw, N. D. (2011). Neural correlates of forward planning in a spatial decision task in humans. *Journal of Neuroscience*, *31*(14), 5526-5539.
- Smid, C. R., Kool, W., Hauser, T. U., & Steinbeis, N. (2020). Model-based decision-making and its metacontrol in childhood. *Preprint at PsyArXiv* <https://doi.org/10.31234/osf.io/ervsb>.
- Sutton, R. S., & Barto, A. G. (1999). Reinforcement learning. *Journal of Cognitive Neuroscience*, *11*(1), 126-134.
- Tyler, M., De George-Walker, L., & Simic, V. (2020). Motivation matters: Older adults and information communication technologies. *Studies in the Education of Adults*, *52*(2), 175-194.
- Van Der Meer, M. A., & Redish, A. D. (2010). Expectancies in decision making, reinforcement learning, and ventral striatum. *Frontiers in neuroscience*, *3*, 1125.
- Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, *20*(3), 174-180.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1870-1884.
- Wasylyshyn, C., Verhaeghen, P., & Sliwinski, M. J. (2011). Aging and task switching: a meta-analysis. *Psychology and aging*, *26*(1), 15.

- West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological bulletin*, 120(2), 272.
- Wilson, A., Fern, A., & Tadepalli, P. (2014). Using trajectory data to improve bayesian optimization for reinforcement learning. *The Journal of Machine Learning Research*, 15(1), 253-282.
- Wyckmans, F., Otto, A. R., Sebold, M., Daw, N., Bechara, A., Saeremans, M., ... & Noël, X. (2019). Reduced model-based decision-making in gambling disorder. *Scientific reports*, 9(1), 19625.
- Zamarian, L., Weiss, E. M., & Delazer, M. (2011). The impact of mild cognitive impairment on decision making in two gambling tasks. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(1), 23-31.
- Ziefle, M., & Bay, S. (2004, September). Mental models of a cellular phone menu. Comparing older and younger novice users. In *International Conference on Mobile Human-Computer Interaction* (pp. 25-37).
- Ziefle, M., & Bay, S. (2006). How to overcome disorientation in mobile phone menus: A comparison of two different types of navigation aids. *Human-Computer Interaction*, 21(4), 393-433.

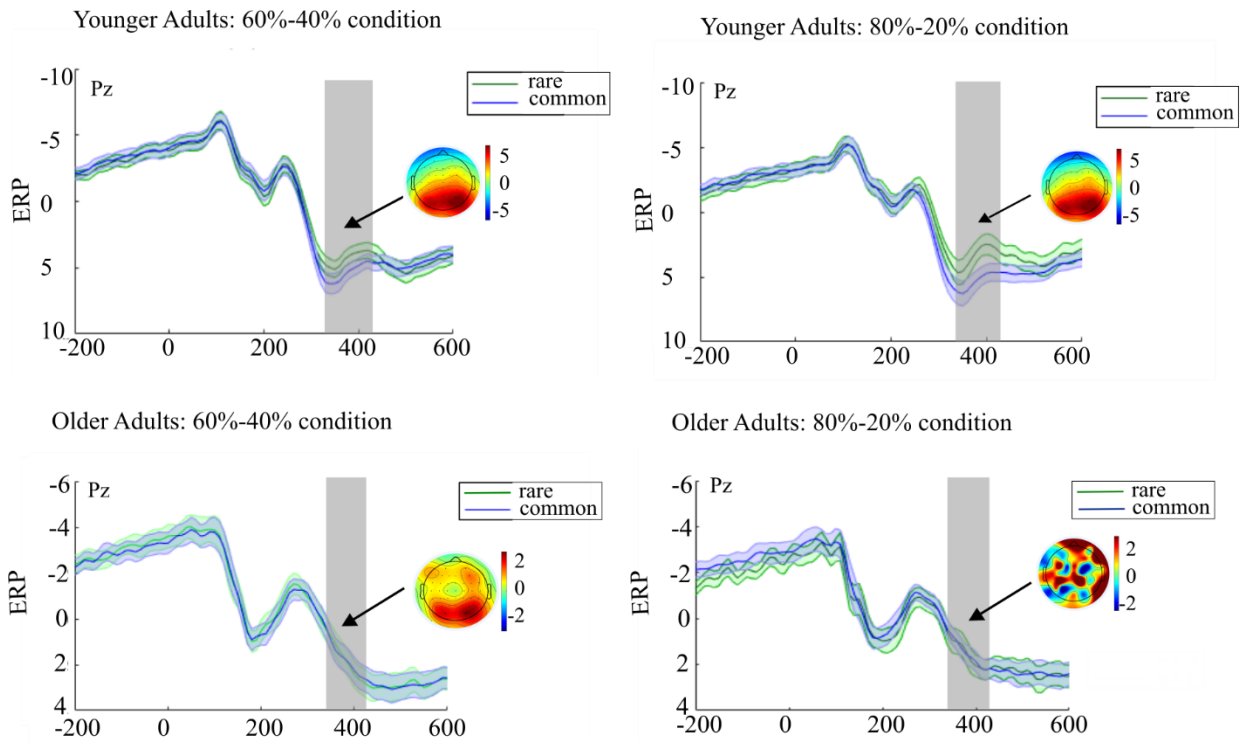
## SUPPLEMENTAL MATERIALS

### APPENDIX A: SUPPLEMENTAL MATERIAL CHAPTER 2

#### *Multiple Single-Trial Robust Regression of stimulus-locked ERPs at the second stage.*

Using transition as the predictor, first stage regressions revealed maximum amplitudes of positive going EEG activity at electrode Pz approximately 370ms after stimulus onset in the 60%-40% condition, and around 380ms after stimulus-onset in the 80%-20% condition, consistent with the P300 component. We therefore used a 330-430ms time window for the second level analyses.

In a separate ANOVA we tested the effects of age group and transition condition on transition related single trial P300 activity. Results revealed a significant main effect of age group ( $F(1,49) = 32.755, p < 0.001$ ) but no significant main effect of transition condition ( $F(1,49) = 3.458, p = 0.066$ )(see fig 5). However, no significant interaction between the factors age group and condition was obtained ( $F(1,49) = 2.183, p = 0.143$ ). These results reflect that the transition predictor explained more variance in the P300 component for the younger adults ( $M = 0.401, SD = 0.926$ ) than the older adults ( $M = -0.426, SD = 0.453$ ). Further, the transition type explained more variance in the P300 in the 80%-20% condition ( $M = 0.135, SD = 1.001$ ) as compared to the 60%-40% condition ( $M = -0.133, SD = 0.623$ ).



*Figure S1.* ERPs elicited by second-stage stimuli at electrode Pz for common (blue) vs rare (green) transitions at the second stage choice period displayed separately for young and older adults and 60%-40% and 80%-20% condition. The shaded areas within each plot represents the window of analysis (330-430ms). Topographies of maximum amplitudes are overlaid on each plot (350ms in the 6040 condition and 410ms in the 8020 condition).

Table S1. Mixed-effects logistic regression on stay probabilities for both age groups.

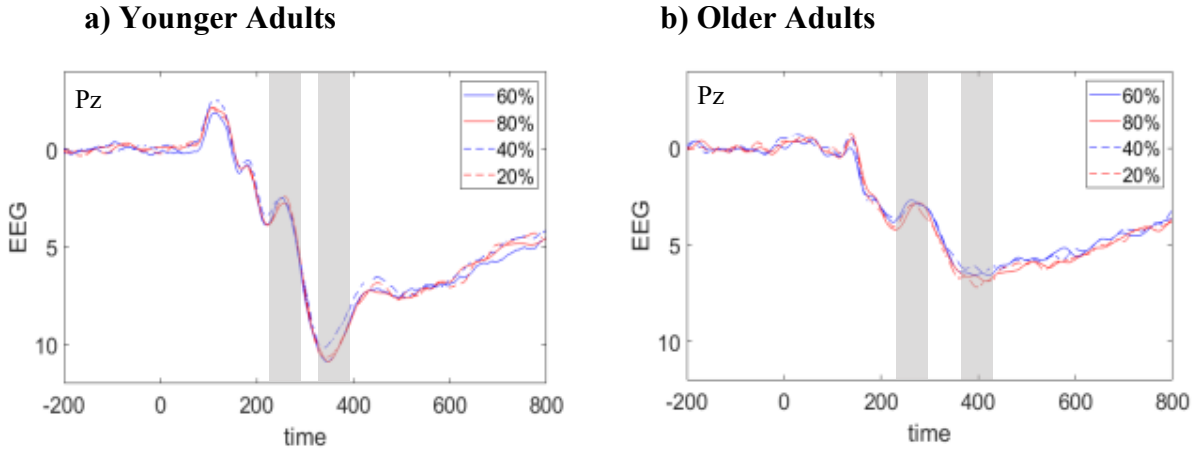
<b>Younger Adults</b>			
<b>Predictor</b>	<b><math>\beta</math></b>	<b><i>SE</i></b>	<b><i>p</i></b>
(Intercept)	1.133	0.100	< 2e-16
Condition	-0.206	0.045	4.97e-06
Transition	0.125	0.040	0.002
Outcome	0.219	0.058	1.720e-04
Condition x Transition	-0.059	0.026	0.021
Condition x Outcome	0.024	0.026	0.358
Transition x Outcome	0.479	0.026	< 2e-16
Condition x Transition x Outcome	-0.215	0.026	< 2e-16
<b>Older Adults</b>			
<b>Predictor</b>	<b><math>\beta</math></b>	<b><i>SE</i></b>	<b><i>p</i></b>
(Intercept)	1.955	0.196	< 2e-16
Condition	-0.002	0.079	0.980
Transition	-0.014	0.040	0.715
Outcome	0.332	0.073	4.94e-06
Condition x Transition	-0.011	0.032	0.745
Condition x Outcome	-0.028	0.033	0.396
Transition x Outcome	0.130	0.032	5.96e-05
Condition x Transition x Outcome	-0.081	0.032	0.010

Table S2. Computational Model: ANOVA Results

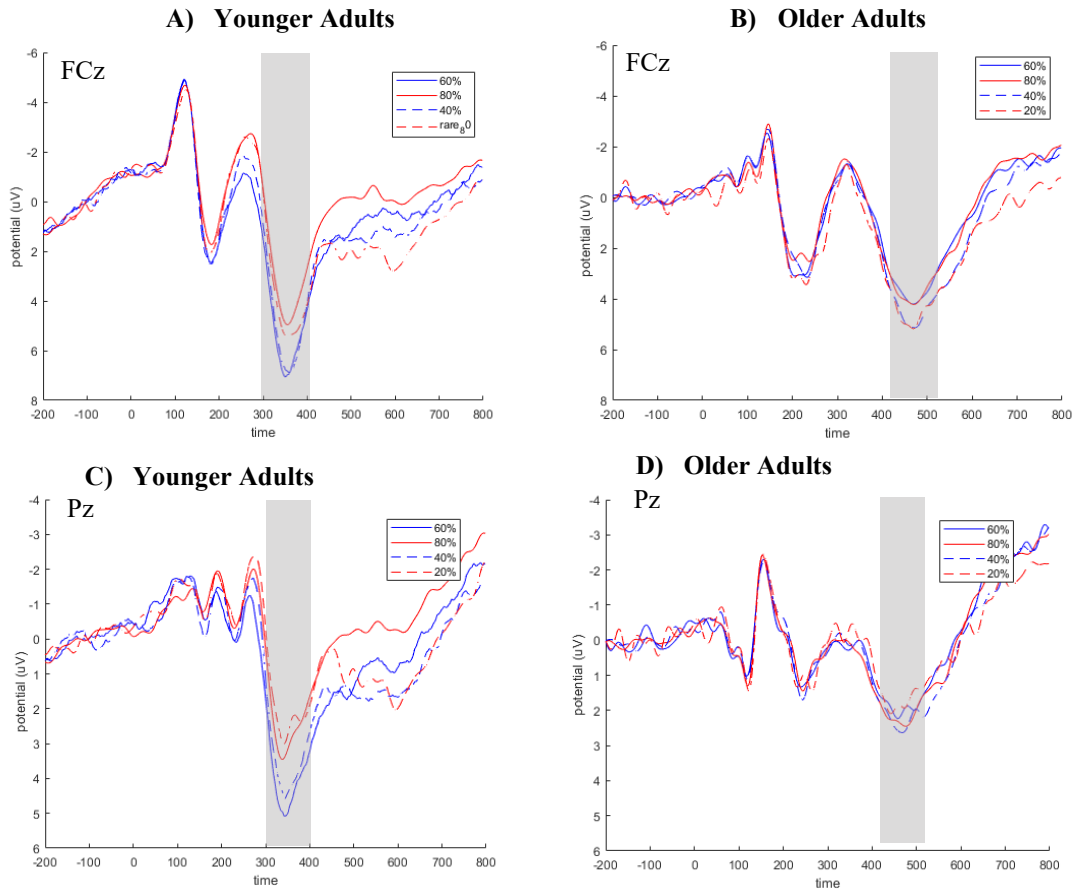
<b>Computational Model Parameter</b>	<b>Predictor</b>	<b>F</b>	<b>p</b>	<b>Partial <math>\eta^2</math></b>
<b><math>\alpha_1</math></b>	Age group	0.088	0.768	0.001
	Condition	1.225	0.274	0.007
	Age group x Condition	6.738	0.012	0.036
<b><math>\alpha_2</math></b>	Age group	4.413	0.041	0.055
	Condition	1.002	0.322	0.007
	Age group x Condition	0.701	0.406	0.005
<b><math>\beta_1</math></b>	Age group	0.022	0.882	< 0.001
	Condition	2.747	0.104	0.021
	Age group x Condition	4.729	0.035	0.036
<b><math>\beta_2</math></b>	Age group	3.347	0.073	0.043
	Condition	0.754	0.389	0.005
	Age group x Condition	0.292	0.591	0.002
<b><math>\pi</math></b>	Age group	17.177	< 0.001	0.197
	Condition	0.347	0.558	0.002
	Age group x Condition	0.530	0.470	0.003
<b><math>\lambda</math></b>	Age group	1.191	0.280	1.132e-02
	Condition	2.763	0.103	2.897e-02
	Age group x Condition	0.001	0.981	5.910e-06
<b><math>\Omega</math></b>	Age group	3.072	0.086	0.040
	Condition	5.288	0.026	0.035

	Age group x Condition	3.355	0.073	0.022
<b>p<sub>common</sub></b>	Age group	4.899	0.032	0.0520
	Condition	7.312	0.009	0.063
	Age group x Condition	8.487	0.005	0.073





*Figure S2.* ERPs elicited by first stage stimuli at electrode Pz, N200 and P300 for **a)** younger adults, and **b)** older adults. These ERPs are displayed separately for the 80%-20% condition (red) and the 60%-40% condition (blue), as well as the common transitions (solid) and the rare transitions (dashed).



*Figure S3.* ERPs elicited by transition phase stimuli at electrode Pz for **a)** younger adults, and **b)** older adults as well as at electrode FCz for **c)** younger and **d)** older adults. These ERPs are displayed separately for the 80%-20% condition (red) and the 60%-40% condition (blue), as well as the common transitions (solid) and the rare transitions (dashed).