

Cortical Grey Matter Mediates Increases in Model-Based Control and Learning from Positive Feedback from Adolescence to Adulthood

Vanessa Scholz,^{1,2} Maria Waltmann,^{1,3} Nadine Herzog,^{3,4} Andrea Reiter,^{1,5} Annette Horstmann,^{3,4,6} and Lorenz Deserno^{1,3,4,7}

¹Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Centre of Mental Health, University of Würzburg, 97080 Würzburg, Germany, ²Donders Institute for Brain, Cognition and Behaviour, Radboud University, 6525 GD Nijmegen, The Netherlands, ³Max Planck Institute for Cognition and Neuroscience, D-04103 Leipzig, Germany, ⁴Integrated Research and Treatment Center AdiposityDiseases, Leipzig University Medical Center, 04103 Leipzig, Germany, ⁵Collaborative Research Center-940 Volition and Cognitive Control, Faculty of Psychology, Technical University Dresden, 01069 Dresden, Germany, ⁶Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, 00014 Helsinki, Finland, and ⁷Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Technical University Dresden, 01069 Dresden, Germany

Cognition and brain structure undergo significant maturation from adolescence into adulthood. Model-based (MB) control is known to increase across development, which is mediated by cognitive abilities. Here, we asked two questions unaddressed in previous developmental studies. First, what are the brain structural correlates of age-related increases in MB control? Second, how are age-related increases in MB control from adolescence to adulthood influenced by motivational context? A human developmental sample ($n = 103$; age, 12–50, male/female, 55:48) completed structural MRI and an established task to capture MB control. The task was modified with respect to outcome valence by including (1) reward and punishment blocks to manipulate the motivational context and (2) an additional choice test to assess learning from positive versus negative feedback. After replicating that an age-dependent increase in MB control is mediated by cognitive abilities, we demonstrate first-time evidence that gray matter density (GMD) in the parietal cortex mediates the increase of MB control with age. Although motivational context did not relate to age-related changes in MB control, learning from positive feedback improved with age. Meanwhile, negative feedback learning showed no age effects. We present a first report that an age-related increase in positive feedback learning was mediated by reduced GMD in the parietal, medial, and dorsolateral prefrontal cortex. Our findings indicate that brain maturation, putatively reflected in lower GMD, in distinct and partially overlapping brain regions could lead to a more efficient brain organization and might thus be a key developmental step toward age-related increases in planning and value-based choice.

Key words: adolescence; development; model-based control; punishment learning; reinforcement learning; reward learning

Significance Statement

Changes in model-based decision-making are paralleled by extensive maturation in cognition and brain structure across development. Still, to date the neuroanatomical underpinnings of these changes remain unclear. Here, we demonstrate for the first time that parietal GMD mediates age-dependent increases in model-based control. Age-related increases in positive feedback learning were mediated by reduced GMD in the parietal, medial, and dorsolateral prefrontal cortex. A manipulation of motivational context did not have an impact on age-related changes in model-based control. These findings highlight that brain maturation in distinct and overlapping cortical regions constitutes a key developmental step toward improved value-based choices.

Received July 21, 2022; revised Dec. 20, 2022; accepted Jan. 13, 2023.

Author contributions: L.D., A.H., N.H. and M.W. designed the study; M.W. and N.H. performed the research; V.S. and L.D. analyzed the data; V.S. and L.D. wrote the manuscript; V.S., A.R., L.D., A.H., N.H. and M.W. edited the manuscript.

This work was directly funded by a grant to L.D. and A.H. from the IFB Adiposity Diseases, Federal Ministry of Education and Research (BMBF), Germany, GN: 01E0150. L.D. also received funding from the German Research Foundation (DFG) as part of the Collaborative Research Centre 265 “Losing and Regaining Control over drug intake” (Project A02), which partially supported this work. A.R. acknowledges support from grants by the German Research Foundation (Deutsche

Forschungsgemeinschaft, DFG RE 4449/1-1, SFB 940-3/B7, RTG-2660/B2) and by a 2020 BBRF Young Investigator Grant. The funders have no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Vanessa Scholz at scholz_v@ukw.de or Lorenz Deserno at deserno_l@ukw.de.
<https://doi.org/10.1523/JNEUROSCI.1418-22.2023>

Copyright © 2023 the authors

Introduction

Value-based learning and decision-making are guided by model-based (MB) and model-free reinforcement learning (RL) systems (Daw et al., 2005, 2011; Daw and Dayan, 2014). The MB system relies on a model of the environment by mapping states, actions, and outcomes in a probabilistic manner (Daw et al., 2005, 2011; Dayan and Niv, 2008; Daw and Dayan, 2014). This enables flexible behavior but is cognitively demanding. MB contributions to control were shown to increase from childhood into adulthood (Decker et al., 2016; Bolenz et al., 2017, 2019; Nussenbaum et al., 2020; Vaghi et al., 2020), which was mediated by cognitive abilities (Potter et al., 2017; Nussenbaum et al., 2020). Our present study replicates these findings and examines hitherto unaddressed research questions. First, what are the neuroanatomical correlates of age-related increases in MB control? Second, are these increases distinctively influenced by motivational context? Finally, does learning from positive and negative feedback change with age?

After a marked increase of gray matter (GM) from infancy to childhood (Knickmeyer et al., 2008; Gilmore et al., 2012), the adolescent brain shows a profound GM reduction in frontal, parietal, and temporal cortices (Blakemore, 2012; Tamnes et al., 2013; Ziegler et al., 2019), potentially from synaptic pruning (Giedd et al., 1999; Gogtay et al., 2004). This may lead to more efficient brain organization, including myelination (Fuhrmann et al., 2015), hypothetically underlying the coinciding improvement in cognition, such as working memory (Bunge and Wright, 2007; Jolles et al., 2011; Tamnes et al., 2013). Previously, MB control was positively related to cognitive abilities including working memory (Eppinger et al., 2013; Otto et al., 2013) and processing speed (Schad et al., 2014; Reiter et al., 2016), and increasing MB control with age was mediated by cognitive abilities (Potter et al., 2017; Nussenbaum et al., 2020). GM in dorsolateral and ventromedial prefrontal cortex correlated with MB control in adults (Deserno et al., 2015; Voon et al., 2015a), and prefrontal and parietal cortices were shown to encode state predictions, a neural signature of MB RL (Gläscher et al., 2010). However, it remains unclear whether structural brain maturation mediates increases in MB control from adolescence to adulthood.

Previous work (Cauffman et al., 2010; Van Leijenhorst et al., 2010; Cohen, 2011; Van Den Bos et al., 2012; Silverman et al., 2015) also suggested that effects of outcome valence on RL (Sutton and Barto, 1998; Daw and Dayan, 2014) may change across development, showing elevated reactivity in adolescents toward rewards overall or relative to punishment (but see Nussenbaum and Hartley, 2019; Rosenbaum et al., 2022). It is feasible that adolescents still differ in their ability to apply and exert the same extent of MB control across different domains given ongoing and protracted brain maturation (Cohen et al., 2016). Also, adolescents may be more strongly affected by the outlook of reward and might use more MB control to gain rewards relative to punishment avoidance. However, the only study examining such effects on MB control found no age dependency of contextual valence on MB control (Bolenz and Eppinger, 2021). Meanwhile, a large body of work links differences in positive versus negative feedback learning to positive and negative reward prediction errors (RPEs; Frank et al., 2004, 2007). Phasic dopamine responses to RPE are asymmetric so that bursts for positive RPEs are larger than dips for negative RPEs (Montague et al., 1996; Schultz et al., 1997). In adolescence, RPE signaling in the ventral striatum is enhanced compared with adults (Cohen et al., 2010). An established test capturing this asymmetry in feedback learning

is derived from a probabilistic selection task (Frank et al., 2004), which has only been used once to study instruction biases across development (Decker et al., 2015).

Thus, we aimed to experimentally separate and study two developmentally relevant aspects of outcome valence, (1) learning in a reward or punishment context and (2) learning from positive versus negative feedback. For this, we recruited a developmental sample that completed structural neuroimaging and a modified 2-step task (Voon et al., 2015b; Doll et al., 2016) to study the structural correlates of the development of MB control for reward and punishment and learning from positive versus negative feedback.

Materials and Methods

Sample

A developmental sample of $n = 103$ participants (age range, 12–50 years, 48 females, 55 males) was recruited as part of a larger study. This subsample was specifically screened to exclude current mental health diagnosis (also see the pre-registered study protocol for more details on the study proceedings and employed material at <https://osf.io/fyn6q>) (Herzog et al., 2023). Participation consisted of two appointments. On day one, each participant completed a modified sequential decision-making task, the 2-step task (Daw et al., 2011) behaviorally, which was followed by a choice test capturing learning from positive versus negative feedback (Frank et al., 2004; Doll et al., 2009, 2011, 2016) after a 30 min break, during which participants either completed another task or questionnaires. This was all part of a larger task battery. They also underwent a battery of cognitive measures [Digit Symbol Substitution Test (DSST) for processing speed (Wechsler, 1997), Digit Span for working memory, (Wechsler, 1997), the Trail Making Test (TMT) for executive functioning including visual attention (Reitan, 1955) and a German vocabulary test assessing verbal intelligence (Schmidt and Metzler, 1992)], also described in more detail in the study protocol at <https://osf.io/fyn6q> (Herzog et al., 2023). On the second day, participants underwent structural MRI. Participants were reimbursed 9 euros per hour for participation and a bonus payment based on task performance. The study was in accordance with the Declaration of Helsinki and approved by the ethics board of the medical faculty at the University of Leipzig (385/17-ek). All participants were informed about study proceedings and gave informed consent before participation.

Sample overview

The following outlines the samples assessed for each analysis. For all samples, the age range was 12–42 years, and additional characteristics are reported accordingly. In the behavioral analysis of MB control in the 2-step task ($n = 101$; male, female, 54:47; mean age = 23.03 years, SD = 7.98), we excluded one participant from the initial sample of 103 participants, who had experienced 140 instead of 200 trials on the task, rendering this participant's learning experience incomparable with the rest of the sample and a second participant who was a considerable age outlier (8 year age gap, age 50).

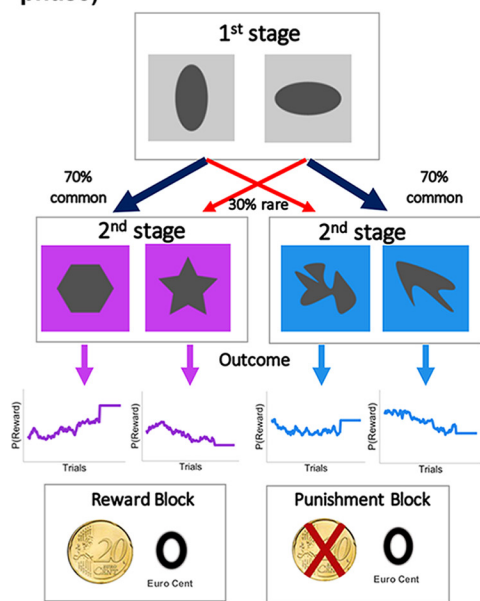
For the behavioral analysis of choice task ($n = 90$; male, female, 52:38; mean age = 23.33, SD = 8.04) fewer participants had completed the task because of technical issues of task presentation ($n = 10$). In line with previous studies (Gillan et al., 2016; Smid et al., 2022), we excluded one more participant with a missing response rate of >95% on the choice task.

For structural brain analysis, $n = 98$ datasets (male, female, 53:45, mean age = 23.14, SD = 8.05) were available, as 3 participants did not undergo scanning. Thus, for the MB measures derived from the 2-step task, brain–behavior correlations could be tested in $n = 98$. Because of the larger dropout for the choice task, brain–behavior correlations could be tested only in $n = 88$; male, female, 51:37; mean age = 23.59, SD = 8.06).

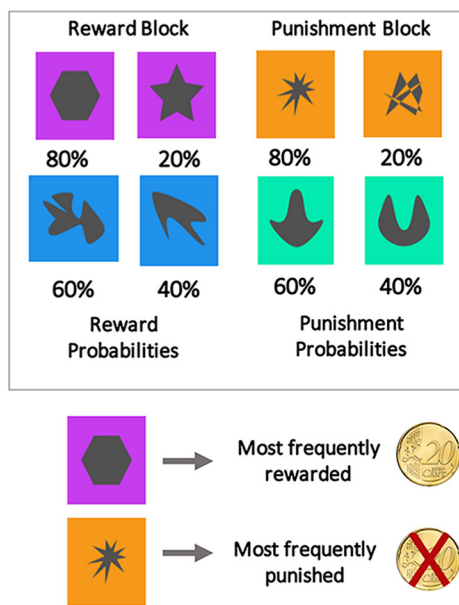
Experimental Tasks

The study included a sequential decision-making task, which encompassed two major modifications to address the research questions outlined in the introduction. These changes included (1) separate reward

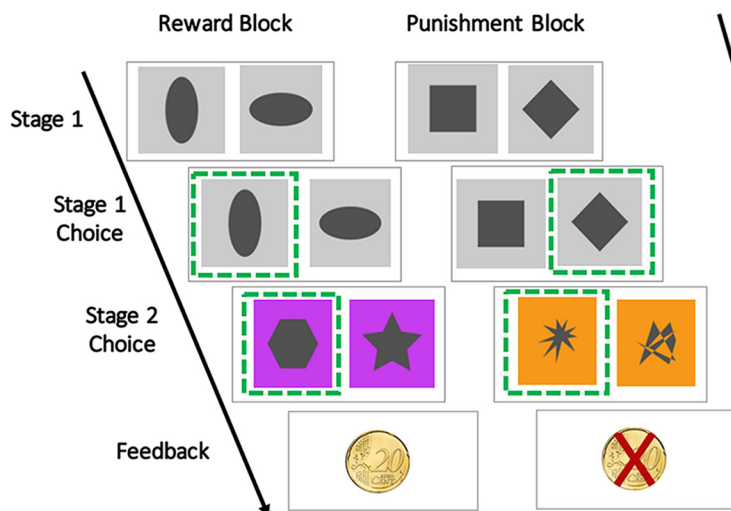
A Modified 2-Step Task (training phase)



B Fixed Outcome Probabilities (last 30 trials)



C Trial sequence training phase



D Test phase (choice without feedback)

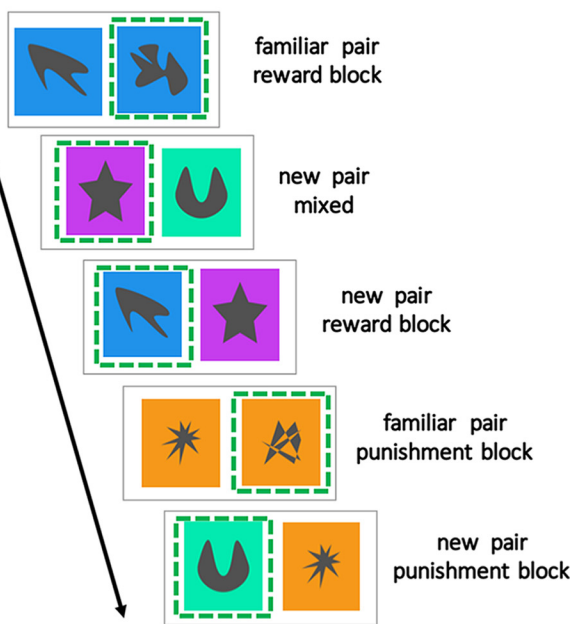


Figure 1. Task set-up sequential decision-making task. **A**, Modified 2-step task: In the sequential decision-making task, a first stage choice led to one of two possible second stages, during which a second choice had to be made. After this second stage choice, participants received a reward or neutral outcome (rewards were replaced by punishments in the punishment context). The probability of receiving a reward or punishment was determined by continuously changing probabilities, i.e. Gaussian random walks. Transition probabilities from stage 1 to stage 2 were fixed and are either considered common (70%) or rare (30%). **B**, Fixed outcome probabilities. Depiction of the fixed outcome probabilities that were assigned to the two familiar cue pairings for the last 30 trials of each motivational context (reward and punishment) in the sequential decision-making task. **C**, Trial sequence sequential decision-making task, presented once for the reward context which employs positive and neutral outcomes, and once for the punishment context. **D**, Overview of choice phase depicting cue selection across five trials without feedback to make sure that no further feedback-based learning occurs. Participants were required to select what they thought was the better cue from either familiar or recombined, new pairs.

and punishment blocks during the task (Voon et al., 2015b) to test effects of motivational context on MB control and (2) a choice test akin to the probabilistic selection task (Frank et al., 2004), following the 2-step task (Doll et al., 2016), testing learning from positive versus negative feedback.

Sequential decision-making task

Similar to Daw et al. (2011), participants were presented with two different cue pairs in the first stage and had to select one cue to continue to

the second stage (Fig. 1). Each cue was associated with a probabilistic transition to one of the two second stage states, a common transition of 70% and a rare transition of 30%. Transition probabilities were fixed across the task. At the second stage, participants again had to choose between two cues and received an outcome (reward, neutral outcome, punishment, depending on the within-subject manipulation of motivational context; see below). Outcome probabilities changed slowly but constantly over time according to a Gaussian random walk. Thus, to maximize outcome in this task, participants needed to track these

continuously changing outcome probabilities. Importantly, participants were explicitly told that the transition structure of the task (common/rare) would be constant over the task and that one transition was going to be more probable than the other. They also underwent practice trials to ensure that they understood the task. By making use of knowledge about the transition structure, individuals could exert MB control over choices. We used a task variant, which included an additional within-subject manipulation of motivational context for positive and negative outcomes (Voon et al., 2015b; Worbe et al., 2016). Each participant completed two task blocks (each 100 trials). The reward block had monetary rewards displayed on the screen (+20 cents) or neutral feedback, whereas the punishment block used punishments (−20 cents) alongside neutral feedback. Block order was randomized across participants to avoid order effects. Participants were instructed to maximize outcome and could do so by learning to select the currently most rewarded cue in the reward block or by avoiding negative outcomes in the punishment block. They were told that they would be paid a bonus dependent on the rewards gained during the experiment.

Choice test

To examine age-dependent differences in learning from positive and negative outcomes and the impact of motivational context on learning, we used a variant of a previously established probabilistic choice task (Frank et al., 2004; Doll et al., 2009, 2011, 2016; Fig. 1D). To enable a choice test subsequent to the 2-step task, reward probabilities remained stable for the last 30 trials in each block (reward/punishment) of the 2-step task. Hence, the previously slowly changing reward probabilities of one second stage pair were fixed to 80% and 20% chance of winning (or losing for the punishment block), whereas the probabilities were fixed at 60% and 40% for the other second stage pair. Thus, in the last 30 trials of each block, participants could learn the cue values in a stable manner (e.g., infer the most frequently rewarded and least frequently punished cues; Fig. 1B). In the ensuing choice task, participants were presented with 28 different cue pairs, which were presented three times amounting to a total of 84 trials. These consisted of four familiar pairs they had previously encountered during the second stage of the 2-step task and 24 unfamiliar pairs from newly combined cues from the second stage. Unfamiliar pairs could be grouped into two categories, (1) two recombined cues derived from the same motivational context (reward or punishment, 8 pairs) and (2) mixed pairs combining cues from the reward and punishment block (16 pairs). Of note, mixed pairs represent a unique feature of this task version and have not been used frequently in previous work (Palminteri et al., 2016). They were introduced to increase the general level of task difficulty and variance of performance thus allowing better assessment of interindividual differences across participants. Participants were instructed to select the best cue from each pair on presentation, and unlike before, they did not receive feedback after having made their choice to disable further feedback-based learning. Thus, the selection task examines the values participants have learned for each cue throughout the previous task phase, that is, the value they encoded for the respective stimulus. Of note, participants were aware of a performance-based bonus payment related to both tasks.

Analysis of 2-step task

To examine which factors had an impact on first-stage choice behavior on the subsequent trial, we computed linear mixed effects models using the lme4 package implemented in R (<http://cran.us.r-project.org>) with the optimizer bobyqa and the maximal number of iterations set to $n = 1e + 8$. The model included participants' trial-by-trial first-stage choices (stay or switch in a given trial n as compared with the previous trial $n-1$) as dependent variable (DV). Second stage feedback (positive vs negative; in the reward block this refers to positive vs neutral outcome, and in the punishment block this refers to neutral vs negative outcome) and transition type (rare vs common) from the previous trial and motivational context (reward vs punishment block) as within-subject fixed effects in the model. The model was estimated with a full random effects structure, that is, all fixed effects were included as random effects as follows:

$$DV_{\text{stay}} \sim \text{feedback} * \text{transition} * \text{valence} + (\text{feedback} * \text{transition} * \text{valence} + 1 | \text{Subject}).$$

To determine whether we could replicate age-dependent changes in MB control as reported by Decker et al. (2016), we extracted the individual slope of the fixed effects interaction term $\text{feedback} * \text{transition}$ as a valence-dependent individual estimate of MB control and correlated it with age. To determine age-dependent changes in MB control as a function of motivational context, we ran the model separately for the reward and punishment blocks and correlated the extracted estimates of $MB\ control_{\text{reward}}$ and $MB\ control_{\text{punishment}}$ with age, respectively. Finally, we also assessed the association between MB control and the quadratic age effect while controlling for the linear effect of age. Given non-normal distributions of some variables, we assessed correlations using Spearman's correlation coefficients.

Cognitive measures and MB control

We also examined the replicability of previous findings (Schad et al., 2014) showing a link between cognition and MB control. For this, we ran mediation analyses using the mediation package (Tingley et al., 2014) implemented in R to test whether the association between age and MB control was mediated by cognitive abilities. We used nonparametric bootstrapping with $n = 10,000$ simulations to determine the average causal mediation effect between age and MB control mediated by cognitive measure.

Analysis of choice test

We studied participants' tendency to learn from positive and negative feedback using mixed-effects modeling and assessed age effects using correlational analysis. Here, we examined the difference between selecting the best cue for reward or punishment pairs (choose 80% rewarded or 20% punished over all cues) relative to avoiding the worst cue in reward or punishment pairs (avoid the cue that was 20% rewarded or 80% punished). This difference captures the shift toward learning more from positive or negative feedback, whereas no difference indicates equal learning from both modalities. We restricted analysis to reward pairs including the most frequent winner (80% positive outcome) or the least frequent winner (20%) and punishment pairs with the least frequent loser (20% negative outcome) or the most frequently punished cue (80%). Selecting the best cue, previously termed cue A, represents an individual's tendency to learn from positive feedback, whereas avoiding the worst cue, also known as B from previous work, is said to capture learning from negative feedback (Frank et al., 2004; Waltz et al., 2007; Doll et al., 2016). The following equation was used to assess the impact of learning from positive and negative feedback on optimal decision-making:

$$DV_{\text{opt}} \sim \text{valence} * AB * + (\text{valence} * AB + 1 | \text{Subject}).$$

Again, we were interested in the association between learning from positive and negative feedback with age (linear and quadratic effect) and examined the correlation coefficient between age and the frequency with which participants selected the best cue or avoided the worst.

Next, we assessed whether each participant's frequency in visiting the second-stage options in the last 30 trials with stable contingencies during the 2-step task affected behavior in the choice task. Therefore, we counted the number of visits of each stimulus during the last 30 trials. Then, for each pair of stimuli presented in the choice task, we included the differences between the number of visits of the two stimuli as covariates to the model. Please note, if each stimulus was visited the same number of times, this difference would be zero. We computed the model as follows:

$$DV_{\text{opt}} \sim \text{valence} * AB * \text{Cue2ndStage} + (\text{valence} * AB + 1 | \text{Subject}).$$

We concluded this analysis by assessing whether MB control was linked to learning from positive or negative feedback in the selection task using correlational analysis.

Structural brain data

Data acquisition structural imaging. Whole-brain T1 weighted image acquisition took place on a 3T Magnetom Skyra scanner (Siemens). Structural data for subsequent voxel-based morphometry analysis was acquired with an echo time (TE), 2.98 ms; repetition time (TR), 2300 ms; voxel size, $1.0 \times 1.0 \times 1.0$ mm; field of view (FOV), 256 mm; 176 slices with a slice thickness of 1 mm. For $n = 16$ participants a different multi-echo sequence with the following parameters was used: TE, 1.96, 5.83, 8.78, 11.73, 15.18 ms; TR, 7000 ms; voxel size, $1.0 \times 1.0 \times 1.0$ mm; FOV, 256 mm; 192 slices with slice thickness of 1 mm.

Preprocessing of structural brain data

To examine morphometrical changes in the gray matter density (GMD) of the brain, structural imaging data were preprocessed and analyzed using the statistical parametric mapping (SPM 12) software (<https://github.com/spm/spm12>) and the Computational Anatomy Toolbox (CAT12; <http://dbm.neuro.uni-jena.de/cat>) version 12.7. Preprocessing followed the default pipeline outlined in the CAT12 manual and encompassed normalization to a template space, tissue segmentation into GM, white matter and CSF, estimation of total intracranial volume (TIV), and smoothing. Smoothing was accomplished using an 8 mm full-width at half-maximum kernel. A 0.1 absolute masking threshold was applied to the data. Before analysis, data were screened, and the weighted average image quality ratings implemented in CAT12 were deemed satisfactory. TIV was then estimated for the entire sample. Evaluating design-orthogonality using SPM 12 provided evidence of nonorthogonality between the TIV regressor and our main effects of interest, namely, MB control and learning from positive feedback. Hence, TIV was not included as a regressor in our GLMs. Instead, we used global scaling to scale the data based on individual TIV values thereby avoiding the removal of variance of interest (<https://neuro-jena.github.io/cat12-help/#intro>). Of note, all GLMs also included an additional regressor of no interest coding for the scanning sequence.

Statistical analysis of GM, MB control and age

We used two GLMs to assess effects of age and MB control on GMD. The first model (GLM_{MBC}) included a regressor with the individual slopes for MB control extracted from the mixed model described above to examine any changes in GMD that were associated with MB control in general. Meanwhile, our second model included regressors for age, MB control, and their interaction (GLM_{MBC_Full}) to determine which changes in GMD that were linked to MB control would remain when controlling for age. The interaction term of this model additionally allowed us to assess whether certain changes in GMD differed as a function of age and MB control. To assess quadratic age effects on GMD, we set up a model including regressors for linear and quadratic age (GLM_{Age2}) in the 2-step sample.

Statistical analysis of GM and feedback learning

To probe the association between GMD and selection task performance, we assessed two more models. The first model consisted of the extent to which participants had learned from positive feedback; that is, they had selected the better cue in a newly combined stimulus pair as results had shown age-dependent changes for positive feedback learning only (GLM_{PosFB}). This was again done to inspect which GMD changes were generally linked to positive feedback learning. The second model had three regressors—Age, A/better cue selected, and the interaction (GLM_{PosFB_Full}). This model allowed us to isolate changes in GMD related to positive feedback learning independent of age effects, to assess age-dependent GMD changes, and to determine whether GMD changes differed as a function of age and feedback learning. We chose to examine positive feedback learning only as positive feedback had shown a strong age effect unlike negative feedback learning. Finally, we again set up one more model to assess quadratic age effects on GMD, namely, GLM_{Age2} including two regressors coding for linear and quadratic age.

Regions of interest

Results of all brain structural analyses were examined (1) using FWE correction of peak levels for multiple comparisons on a whole-brain level and (2) using small-volume correction relying on three a priori defined

masks that were derived from Automated Anatomic Labeling software (Tzourio-Mazoyer et al., 2002). This included a ventromedial prefrontal cortex (vmPFC) mask comprising the superior medial frontal and medial orbital gyrus, a dorsolateral PFC (dlPFC) mask based on the middle frontal gyrus, and the parietal cortex (inferior parietal and angular gyrus), which was motivated by previous accords of the involvement of those regions in MB control in fMRI studies using a similar task (Gläscher et al., 2010; Daw et al., 2011; Deserno et al., 2015; Voon et al., 2015a). Results were considered significant with a p value < 0.02 (0.05/3 for three regions of interest) to correct for multiple comparisons.

Mediation analysis

Given the lack of previous work examining the structural correlates of age-dependent changes of MB control and feedback learning, we relied on mediation analysis to examine whether maturational changes in GMD mediated the association between age and positive feedback learning. For this, we used age as independent variable (IV), MB control (or positive feedback learning) as DV, and GMD as mediator.

To assess whether certain GMD changes mediated the association between age and MB control or age and learning from positive feedback, we used the previously computed GLM with a regressor for MB control and a second GLM including age as regressor. For both GLMs, we created an F -contrast assessing any associated changes in either direction for the regressor age and MB control, respectively. Each statistical map was thresholded at $p = 0.001$ and cluster size ≥ 20 voxels at the whole-brain level and subsequently exported as a mask image. We then created conjunction masks from the two F -contrast masks and the parietal cortex mask, dlPFC and mPFC masks, respectively. We then extracted GMD estimates from these mask regions using the `get_totals` script by Ged Ridgway (http://www.cs.ucl.ac.uk/staff/G.Ridgway/vbm/get_totals.m). Of note, as no clusters in the vmPFC were significantly associated with MB control, a conjunction mask did not yield a common region from which to extract GMD estimates. This meant we could not assess the overall mediation effect across all regions but only across two regions before examining the individual mediation effects of the ROIs (here for parietal cortex and dlPFC). For this we created an average score of the scaled GMD estimates of both ROIs. For mediation analysis, we relied on the mediation package in R.

To follow up on age-dependent effects for feedback learning, we used the same approach as outlined for the mediation analysis including MB control to analyze mediation effects related to feedback learning. The only difference was that we now based the conjunction mask on the GLM for (positive) feedback learning. Of note, as we only found age-dependent effects of learning from positive feedback, this analysis focused on positive feedback learning. Again, we examined the mediating role of GMD derived from the vmPFC, dlPFC, and parietal cortex conjunction masks. For this, we first ran a mediation analysis assessing the overall effects of GMD across all ROIs on the association between age and positive feedback by computing a GMD average score and subsequently assessed the impact of those three regions separately.

Data availability

Anonymized data and analysis code for this study can be found at <https://osf.io/7zw62/>.

Results

Replication: MB control increases with age ($n = 101$)

Using correlational analysis, as done by Decker et al. (2016), we replicated age-dependent increases of MB control [$r_s(99) = 0.31$, p value = 0.002; Decker et al., 2016; Nussenbaum et al., 2020; Fig. 2]. Also, MB control in the punishment and reward block both significantly and comparably correlated with age [MB control_{Punishment}, $r_s(99) = 0.28$, $p = 0.004$; MB control_{Reward}, $r_s(99) = 0.27$, $p = 0.007$] indicating that age-dependent improvements in MB control did not differ as a function of motivational context (reward gain vs punishment avoidance blocks). This replicates the previously reported absence of general effects of motivational context (Bolenz and Eppinger, 2021). MB control did not correlate with quadratic

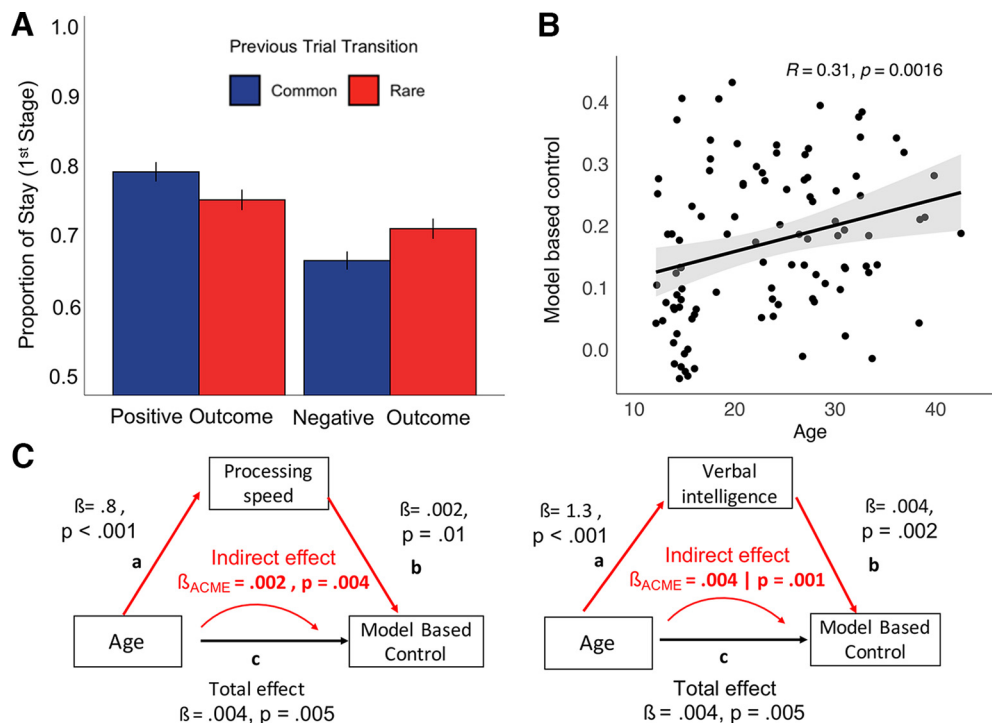


Figure 2. **A**, Depiction of task performance overall. The proportion of times participants stayed with their previous choice on the first stage is depicted as a function of outcome (positive versus negative) and transition type (common versus rare) encountered on the previous trial. The error bars represent ± 1 standard error of the mean. **B**, The scatterplot shows the association between model-based control with age and includes the best-fitted regression line. The shaded area represents the 95% confidence interval. **C**, Mediation effects of the relationship between age and model-based control through cognitive abilities, namely processing speed as measured by the DSST and verbal intelligence indexed through WST scores. *P* values below 0.05 are considered significant.

age effects when controlling for the linear age effect [$r_s(99) = 0.2$, *p* value = 0.1].

Replication: cognition mediates age-related increases in MB control

Mediation analysis provided evidence of a significant partial mediation of processing speed measured using the DSST, accounting for 42.1% ($p = 0.006$) of the total effect of age on MB control [indirect effect = 0.002, CI (0.0005–0.003), $p = 0.004$; direct effect = 0.003, CI (–0.0007–0.01), $p = 0.1$; total effect = 0.004, CI (0.002–0.01), $p = 0.006$]. For verbal intelligence (WST = Wortschatztest), mediation analysis suggested a full mediation for the effect of age on MB control [$p = 0.005$; indirect effect = 0.005, CI (0.002–0.01), $p = 0.002$; direct effect = –0.0001, CI (–0.004–0.002), p value = 0.98; total effect = 0.004, CI (0.002–0.01), $p = 0.003$; Fig. 2]. Both effects were significant after multiple comparisons correction ($\alpha = 0.05/6 = 0.008$). No significant mediation effects were observed for working memory, short-term memory capacity, visual attention, and general executive functioning (all p values > 0.05).

Structural brain correlates of MB control ($n = 98$)

For the GLM_{MBC} , including individual estimates of MB control and a covariate controlling for the structural sequence, no significant association for MB control with GMD survived FWE correction on the whole-brain level. Using small volume correction (SVC) in our a priori defined regions of interest, we found GMD in the parietal cortex to be associated with MB control (MNI peak coordinates, –48, –50, 57; $k = 494, t = 4.42, p_{FWE \text{ Peak corr}} = 0.008$), which did not reach significance for the effect of MB control in dlPFC GMD (MNI peak coordinates, –32, 57, 2; $k = 10, t = 3.29, p_{FWE \text{ Peak corr}} = 0.3$; Fig. 4A). In the vmPFC, no suprathreshold clusters were identified. In the GLM_{MBC_Full} ,

age negatively correlated with GMD across the entire cortex, most prominently for a large cluster comprising the right frontal superior gyrus, the left temporal middle gyrus, and the right supramarginal gyrus (MNI peak coordinates, 18, 56, 16; $k = 37120, t = 6.02, p_{FWE \text{ Peak corr}} = 0.001$; Fig. 4B). We did not find significant association with GMD for MB control while controlling for the effects of age nor for the interaction for age and MB control in this whole-brain analysis ($p_{FWE \text{ corr peak/cluster}} > 0.05$). Likewise, we did not find any significant clusters using SVC for our three predefined regions for the MB control regressor or the interaction regressor. Assessment of the whole-brain effects of GLM_{Age^2} did not yield any significant effects of the quadratic age term (all $p_{FWE \text{ Peak corr}} > 0.05$), so that we consequently refrained from assessing additional models including quadratic age effects.

Next, we examined whether overall GMD across parietal regions and dlPFC mediated the association between age and MB control. The overall mediation model showed a trend, indicating that overall GMD across both regions accounted for 45% of the relationship between age and MB control ($p = 0.054$). We followed this up with individual mediation models to determine whether one of the regions might have a stronger impact. In an individual mediation model, the association between age and MB control was partially mediated by GMD in the parietal cortex, with GMD accounting for 67.7% of the total effect of the relationship between age and MB control ($p = 0.007$; Fig. 4D). Meanwhile, the mediation effect of GMD in the dlPFC was nonsignificant ($p = 0.7$). Given the absence of significant quadratic age effects on MBC or GMD, we abstained from assessing the mediation effects of GMD on the relationship between age² and MB control.

Learning from positive and negative feedback ($n = 90$)

Examining the ability to learn from positive and negative feedback using mixed-effect models, we found a significant main

effect of learning from positive and negative feedback ($\beta = 0.361$, $SE = 0.1$, $\chi^2 = 12.13$, $p < 0.001$), indicating that overall participants were better in choosing the best cue (learning from positive feedback) relative to avoiding the worst cue (Select best: Mean = 68.7% (SD = 23.2) vs Avoid Worst: Mean = 57.8% (SD = 23.2); Fig. 3A). Using correlational analysis, we found that learning from positive feedback improved with age [$r_s(88) = 0.35$, $p < 0.001$], whereas the ability to learn from negative feedback did not [$r_s(88) = 0.07$, $p = 0.5$; compare Fig. 3B,C]. Meanwhile, assessing the partial correlation between the quadratic age effect and the ability to learn from positive or negative feedback while controlling for linear age effects showed no significant effects [Select Best, $r_s(88) = -0.12$, $p = 0.3$; Avoid Worst, $r_s(88) = -0.20$, $p = 0.1$]. We did not find evidence that the differences between the number of visits of the two stimuli significantly affected the choice behavior individuals showed on this task, indicated by the absence of a significant main effect ($\beta = -0.010$, $SE = 0.1$, $\chi^2 = 0.2$, $p = 0.6$) or any significant two-way (Cue2ndStage \times AB, $\beta = 0.046$, $SE = 0.1$, $\chi^2 = 0.7$, $p = 0.4$; Cue2ndStage \times valence, $\beta = -0.013$, $SE = 0.1$, $\chi^2 = 0.2$, $p = 0.7$) or three-way interactions (Cue2ndStage \times valence \times AB, $\beta = 0.103$, $SE = 0.1$, $\chi^2 = 1.6$, $p = 0.2$).

MB control and feedback learning ($n = 90$)

MB control was positively correlated with positive feedback learning in the choice task [$r_s(88) = 0.28$, $p = 0.007$]. Negative feedback learning did not correlate significantly with MB control [$r_s(88) = 0.15$, $p = 0.2$; compare Fig. 3D].

Brain structural correlates of positive feedback learning ($n = 88$)

We then examined the neuroanatomical correlates of age-dependent increases in learning from positive feedback. In the GLM_{PosFB} , we found positive feedback learning to be negatively related to GMD in two clusters on the whole-brain level. Specifically, in the left and right frontal superior medial gyrus (vmPFC, MNI peak coordinates, 9, 69, 22; $k = 1037$, $t = 5.02$, $p_{FWE\ corr\ peak} = 0.02$) and to GMD in a cluster in the right inferior and middle temporal gyrus (MNI peak coordinates, 63, -60, -12; $k = 446$, $T = 5.08$, $p_{FWE\ corr\ peak} = 0.02$; Fig. 5A). Using SVC, we also found a significant cluster in the parietal ROI (MNI peak coordinates, -45, 63, 24; $k = 71$, $t = 4.12$, $p_{FWE\ corr\ peak} = 0.02$) and a significant cluster in the dlPFC ROI (MNI peak coordinates, 20, 63, 26; $k = 104$, $T = 4.30$, $p_{FWE\ corr\ peak} = 0.02$). In the GLM_{PosFB_Full} we saw widespread GMD changes that were negatively associated with age. For positive feedback learning, only one cluster in the supplemental motor area survived whole-brain correction on the peak level (MNI peak coordinates, -3, 16, 70; $k = 390$, $t = 4.97$, $p_{FWE\ corr\ peak} = 0.03$), whereas the previously significant frontal cluster from the first GLM failed to reach significance when controlling for age effects. We found significant clusters in the

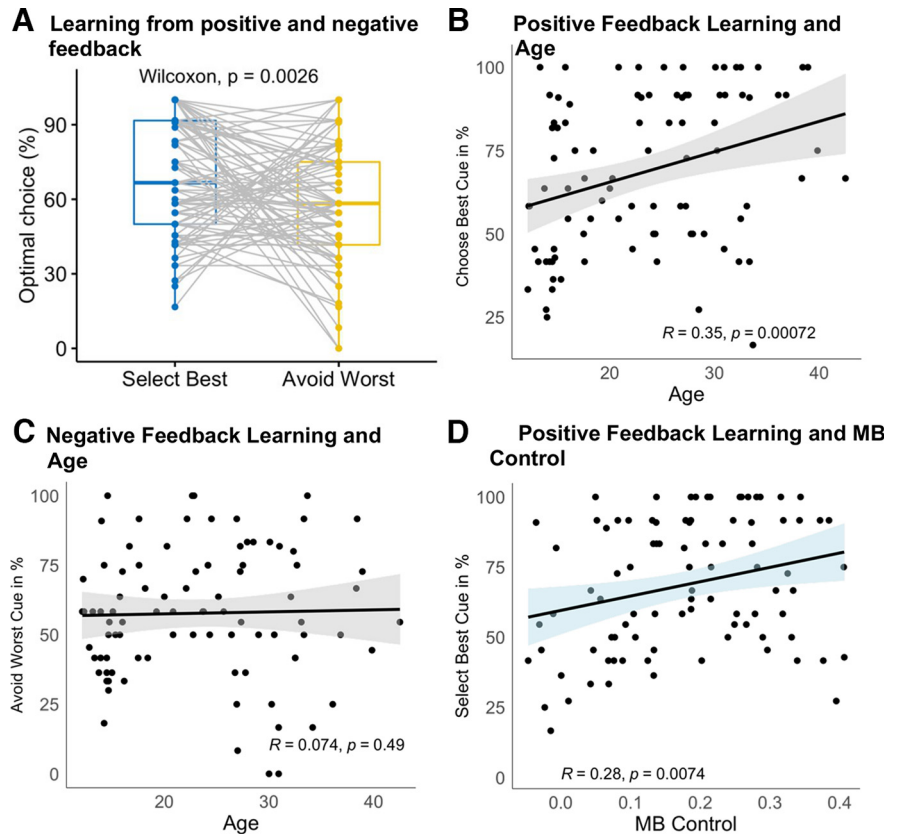


Figure 3. *A*, Overview of participants' performance when choosing the best cue and avoiding the worst cue across reward and punishment block. Participants show better performance when choosing the best cue, i.e. they learn better from positive feedback relative to learning from negative feedback or avoiding the worst cue across blocks. *B*, The scatterplot depicts the age dependent increase of learning from positive feedback. *C*, Learning from negative feedback did not show age effects. *D*, MB control was positively associated with positive feedback learning as can be seen in the scatterplot. The shaded areas in Panel *B–D* all depict the 95% confidence interval and all scatterplots include the best-fitted regression line.

parietal ROI (MNI peak coordinates, 60, -44, 45; $k = 197$, $t = 4.21$, $p_{FWE\ corr\ peak} = 0.02$) and dlPFC (MNI peak coordinates, -33, 62, 8; $k = 546$, $t = 4.72$, $p_{FWE\ corr\ peak} = 0.005$) using SVC, whereas clusters in the vmPFC did not survive multiple comparisons correction. GMD changes were not significantly associated with the regressor for age and positive feedback learning.

When assessing the overall mediation effect of GMD across all regions on the association between age and positive feedback learning, overall GMD accounted for up to 60.5% of the total effect of the relationship of age and positive feedback learning ($p = 0.003$). In the individual mediation models, we found the association for age and learning from positive outcomes to be partially mediated by GMD in all three regions of interest. GMD in the vmPFC accounted for up to 31.2% of the total effect of the relationship of age and positive feedback learning ($p = 0.046$), whereas dlPFC GMD also mediated this association, explaining 42.4% ($p = 0.01$), as did GMD in the parietal cortex, explaining up to 57.9% ($p = 0.01$) of the effect (Fig. 5B–D). Again, we refrained from assessing GMD mediation effects on the association between quadratic age effects and feedback learning given the lack of significant results.

Discussion

We found changes in the association between age and MB control and positive feedback learning to be mediated by reduced parietal GMD. Moreover, the association between age and

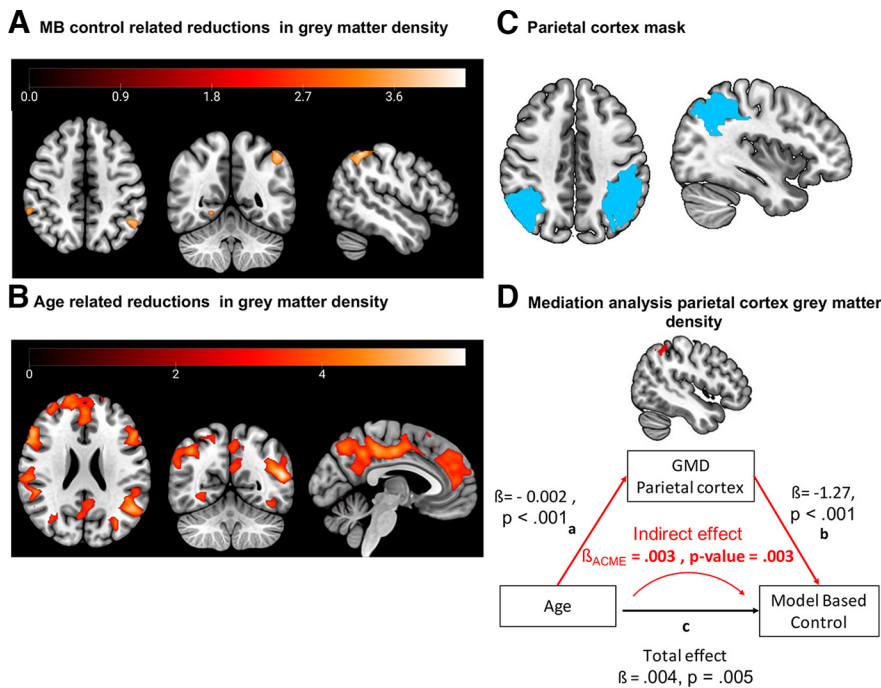


Figure 4. *A*, GMD reductions in the parietal cortex (SVC: MNI Peak coordinate: $-48 -50 57$, $k = 494$, $t = 4.42$, $p_{FWE\ Peak\ corr} = .008$) related to MB control in GLM_{MBC} . *B*, Age related GMD reductions found in the $GLM_{MBC\ Full}$ on the whole-brain level, with the strongest association found for a cluster entailing the right frontal superior gyrus, left temporal middle gyrus and right supramarginal gyrus (MNI Peak coordinate: $18\ 56\ 16$, $k = 37120$, $t = 6.02$, $p_{FWE\ Peak\ corr} = .001$). For visualization purposes, the brain map was thresholded at voxel > 1099 depicting clusters significant at the whole brain level. *C*, Depiction of the parietal cortex mask derived from the AAL atlas incorporated in the WFU PickAtlas Tool v3 (https://www.nitrc.org/projects/wfu_pickatlas/). *D*, Mediation analysis showing the mediation effect of GMD extracted from an *a-priori* defined brain mask in the parietal cortex on the association between age and MB control.

positive feedback learning was partially explained by differences in GMD in the vmPFC and dlPFC. Thus, we provide a potential neuroanatomical correlate for previous findings of age-dependent increases of MB control, which were mediated by cognitive abilities. However, we found no evidence for motivational context relating to age-dependent changes in MB control. Finally, individuals relying more on MB control were more adept at learning from positive feedback, suggesting improved value-based choices. Our findings thus provide significant new insight into potential neuroanatomical correlates of age-related increases in MB control and positive feedback learning while highlighting the need for considering motivational context and outcome valence within a developmental research framework.

To our knowledge, this is the first study examining the neuroanatomical correlates of developmental changes in MB control in different motivational contexts. Our finding of parietal GMD mediating a considerable portion ($>60\%$) of the relationship between age and MB control corresponds to functional imaging work showing this region to encode the neural signature of MB control, a state prediction error, computing the difference between current state and observed state transition (Gläscher et al., 2010). It also agrees with the suggested role of posterior parietal regions as an integration hub of spatial, temporal, and reward information (Roitman et al., 2007).

We also find GMD in the parietal cortex, vmPFC, and dlPFC to partially mediate the association between age and positive feedback learning. The parietal cortex has been implicated in mapping ordinal relationships among cues, a highly relevant feature when learning the relative rank order of task cues (Munoz et al., 2020). The relevance of vmPFC for feedback learning is also

supported by work tying mPFC and mOFC to adaptive decision-making and subjective value encoding (Rushworth et al., 2011) for positive (Kringelbach, 2005) and negative outcomes (Tom et al., 2007). For the dlPFC, previous literature extensively highlighted its role in working memory processes (Rottschy et al., 2012; Nee et al., 2013) further supported by lesion studies in humans (Barbey et al., 2013) and nonhuman primates (Butters and Pandya, 1969; Butters et al., 1971; Levy and Goldman-Rakic, 1999). Interestingly, one study showed a modulatory effect on reward and punishment sensitivity in a probabilistic RL task during noninvasive dlPFC stimulation, possibly mediated through increased dopaminergic release in the ventral striatum (Ott et al., 2011). Thus, it is conceivable that the reported mediation effects of dlPFC GMD on the association between age and positive feedback learning is conveyed through working memory processes facilitated in the dlPFC, which get updated in a feedback-dependent manner.

Assessment of the overall mediation effect of GMD suggests that instead of a unique contribution of distinct brain regions, age-associated changes in feedback learning likely emerge from changes in a frontoparietal network, which has been repeatedly linked to complex, higher-

level cognitive functions such as cognitive control (Niendam et al., 2012; Cocuzza et al., 2020). This notion is further supported by longitudinal work indicating task performance during feedback learning to be linked to functional activity in a frontoparietal network, also showing age-dependent activity patterns (Peters et al., 2016).

Previous developmental studies consistently reported age-dependent increases in MB control (Decker et al., 2016; Bolenz et al., 2017; Nussenbaum et al., 2020; Vaghi et al., 2020), but they were not designed to disentangle motivational context effects. Interestingly, we see a lack of motivational context effects on MB control (and age-dependent effects therein), despite ample empirical evidence describing contextual effects on RL (Louie and De Martino, 2014; Palminteri et al., 2015; Bavard et al., 2018; Pischetta et al., 2020; Palminteri and Lebreton, 2021). This replicates a previous null finding of another study also examining motivational context effects in a developmental sample (Bolenz and Eppinger, 2021). This absence might be explained by the way motivational context can distinctively have an impact on value updating depending on the reference point (Palminteri et al., 2015; Palminteri and Lebreton, 2021), potentially facilitated by our block design. For illustration, in a punishment context, successful punishment avoidance (neutral feedback) can be perceived as rewarding, reinforcing the chosen option, whereas in a reward context, neutral feedback can be punitive, given the potential of gaining reward (Palminteri et al., 2015). It is thus conceivable that participants adjusted their reference point according to motivational context. Another explanation might arise from the used task version as more MB control does not translate into increased payouts or loss avoidance (Kool et al.,

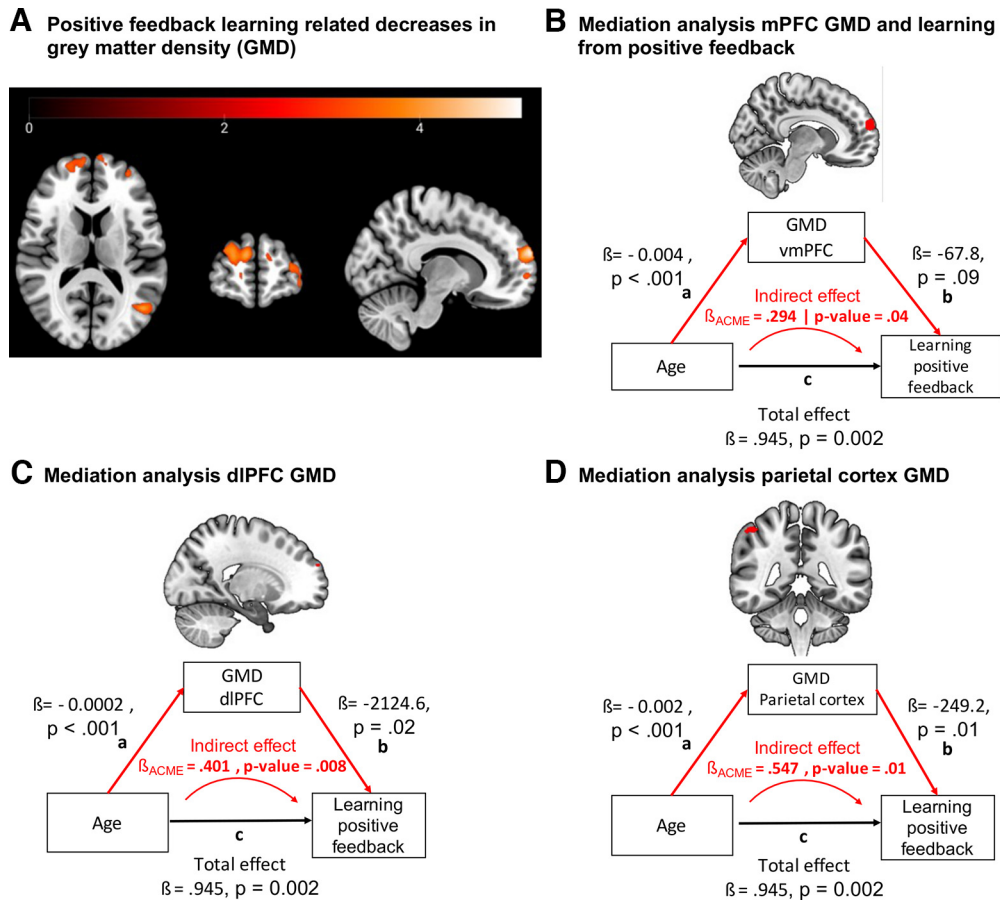


Figure 5. *A*, GMD reductions in the left and right frontal superior medial gyrus (MNI peak coordinates, 9, 69, 22, $k = 1037$, $T = 5.02$, $p_{FWE\ corr\ peak} = 0.02$) and in the right inferior and middle temporal gyrus (MNI peak coordinates, 63, -60 , -12 , $k = 446$, $T = 5.08$, $p_{FWE\ corr\ peak} = 0.02$) were significantly associated with increased learning from positive feedback on the whole-brain level as reported for the first GLM including MB control only as regressor. *B–D*, Structural mediation analysis. Depiction of significant mediation effect of GMD in a cluster in the mPFC, dlPFC, and parietal cortex on the relationship between age and learning from positive feedback; p values below 0.05 are considered significant.

2016), potentially resulting in reduced sensitivity for motivational context. Still, this appears unlikely given previous work using a task variant addressing this short-coming, which also failed to report motivational context effects on MB control (Bolenz and Eppinger, 2021). Thus, while replicating previous findings of age-dependent increases of MB-control (Decker et al., 2016; Nussenbaum et al., 2020), we find no support for our hypothesis that adolescents exert relatively more effortful control by using more MB control in a reward versus punishment context.

Choice test performance indicated an age-dependent increase for positive feedback learning, whereas negative feedback learning seemed stable across development. Previous studies have already pointed to adults exhibiting better reward relative to punishment learning compared with adolescents (Van Der Schaaf et al., 2011; Palminteri et al., 2016), and increasing reward learning rates from childhood into adulthood suggest faster adaptation following rewards (Van Den Bos et al., 2012). Also, Decker et al. (2015) showed developmental changes using a modified version of this choice task, with children and adolescents displaying a less optimal choice pattern relative to adults. In contrast, other studies reported age-dependent increases in negative learning rates, whereas positive learning rates remained stable (Pauli et al., 2022; Rosenbaum et al., 2022). Adolescents focused more on worse-than-expected outcomes during learning, and this tendency was associated with a memory recall bias (Rosenbaum et al., 2022). Interestingly, it has been proposed that

the often-reported elevated reward responsiveness in adolescents might not originate from enhanced reward learning but instead reflect a tendency of more pronounced action initiation (Pauli et al., 2022). Heterogenous findings might result from significant differences in task design but need further investigation to better understand the dynamics of age-dependent valence asymmetries in RL (Nussenbaum and Hartley, 2019). Overall, our results show that although recruitment of MB control appears unaffected by reward or punishment feedback, the way we learn and encode outcome-representations seems biased towards positive feedback learning.

Interestingly, task performance across tasks was connected, with more MB control linked to better positive feedback learning. This suggests that independent of motivational context, more MB control facilitates improved encoding of outcome-value representations for the respective cues in the first task phase allowing for better performance when selecting the better cue during the choice task. This did not apply to negative feedback learning. Given that parietal regions were implicated in mediating both the association between MB control and positive feedback learning with age, maturation of this brain region might underly further development of key processes such as value or cue representation relevant for both constructs. In mice, silencing the activity of the posterior parietal cortex (PPC) leads to considerable impairment of categorizing newly encountered stimuli and cue recategorization, thus suggesting a pivotal role of the parietal cortex in cue representation and fundamental

mechanisms underlying basic learning processes (Zhong et al., 2019). Further support for the notion that the parietal cortex supports key processes of choice behavior and action selection comes from another animal study showing reduced effects of previous choice history on subsequent action selection after temporary optogenetic deactivation of the PPC (Hwang et al., 2017).

Importantly, although the choice task was designed to distinguish model-free (MF) and model-based components, it has been posited that the asymmetric, valenced choice signature of feedback learning could be model free (Frank, 2005; Collins and Frank, 2014; Doll et al., 2016). Thus, it might be surprising that the latter signature was linked positively to MB control, at least when adhering to a strict dual-system separation. Still, previous work indicated that such a distinction might not be appropriate by showing that the MF system can access MB information (Daw et al., 2011; Moran et al., 2019; Deserno et al., 2021). Second, the association might arise from inherent experimental features, given that our MB index and the choice signature of feedback learning were derived from two highly interdependent tasks. Finally, the positive correlation between both variables might reflect the ability to sharply represent values of distinct choice options, a feature supportive for both MF and MB processes. Future studies are thus warranted to precisely characterize these task signatures.

Our data provide new insights into how brain structural changes have an impact on the age-dependent maturation of decision processes. Unlike previous developmental studies, which primarily compared children and adolescents with adults under 30 years (Van Den Bos et al., 2012; Decker et al., 2016; Potter et al., 2017), our study fills an important gap in the literature by assessing individuals across a broader age range including middle adulthood. Our findings of increased MB control and feedback learning provide exciting new insights into decision and learning systems, both of which might have the capacity of further improving into middle adulthood. Interestingly, older individuals learned better from negative feedback relative to younger adults (Eppinger and Kray, 2011), suggesting that feedback learning in old age might not follow the trajectory seen in adolescence and adults (Singh-Manoux et al., 2012; Ferreira et al., 2015).

Regarding limitations, first, our study has a cross-sectional design, and our findings are correlational in nature. More longitudinal studies (Vaghi et al., 2020) are needed to replicate our findings while tracking interindividual developmental trajectories of the reported associations, which can otherwise go unnoticed in cross-sectional designs (Koolschijn et al., 2011). This seems especially important considering substantial variations in structural brain development across subcortical and cortical regions (Wierenga et al., 2014; Mills et al., 2021). Second, the use of different imaging sequences is not desirable but was controlled for.

In sum, our study shows that age-related changes of MB control and positive feedback learning are mediated by GMD in the parietal cortex, dlPFC, and vmPFC. Meanwhile, more MB control also translates into improved selection of cues linked to better outcomes. Our results underline the importance of taking into account valence effects by teasing apart the role of (1) motivational context and (2) learning from positive and negative outcomes when studying decision and learning processes in the framework of developmental research. The potential implications of our findings for neurodevelopmental disorders such as ADHD and compulsivity, including how developmental trajectories of MB control and feedback learning might diverge in non-healthy development, should be the focus of future longitudinal studies.

References

- Barbey AK, Koenigs M, Grafman J (2013) Dorsolateral prefrontal contributions to human working memory. *Cortex* 49:1195–1205.
- Bavard S, Lebreton M, Khamassi M, Coricelli G, Palminteri S (2018) Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nat Commun* 9:4503.
- Blakemore SJ (2012) Imaging brain development: the adolescent brain. *Neuroimage* 61:397–406.
- Bolenz F, Eppinger B (2021) Valence bias in metacontrol of decision making in adolescents and young adults. *Child Dev* 93:e103–e116.
- Bolenz F, Reiter AMF, Eppinger B (2017) Developmental changes in learning: computational mechanisms and social influences. *Front Psychol* 8:2048.
- Bolenz F, Kool W, Reiter AM, Eppinger B (2019) Metacontrol of decision-making strategies in human aging. *Elife* 8:e49154.
- Bunge SA, Wright SB (2007) Neurodevelopmental changes in working memory and cognitive control. *Curr Opin Neurobiol* 17:243–250.
- Butters N, Pandya D (1969) Retention of delayed-alternation: effect of selective lesions of sulcus principalis. *Science* 165:1271–1273.
- Butters N, Pandya D, Sanders K, Dye P (1971) Behavioral deficits in monkeys after selective lesions within the middle third of sulcus principalis. *J Comp Physiol Psychol* 76:8–14.
- Cauffman E, Shulman EP, Steinberg L, Claus E, Banich MT, Graham S, Woolard J (2010) Age differences in affective decision making as indexed by performance on the Iowa gambling task. *Dev Psychol* 46:193–207.
- Cocuzza CV, Ito T, Schultz D, Bassett DS, Cole MW (2020) Flexible coordinator and switcher hubs for adaptive task control. *J Neurosci* 40:6949–6968.
- Cohen JR, Asarnow RF, Sabb FW, Bilder RM, Bookheimer SY, Knowlton BJ, Poldrack RA (2010) A unique adolescent response to reward prediction errors. *Nat Neurosci* 13:669–671.
- Cohen AO, Breiner K, Steinberg L, Bonnie RJ, Scott ES, Taylor-Thompson KA, Rudolph MD, Chein J, Richeson JA, Heller AS, Silverman MR, Dellarco DV, Fair DA, Galván A, Casey BJ (2016) When is an adolescent an adult? Assessing cognitive control in emotional and nonemotional contexts. *Psychol Sci* 27:549–562.
- Cohen M (2011) Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *Neuroimage* 55:1373–1383.
- Collins AGE, Frank MJ (2014) Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol Rev* 121:337–366.
- Daw ND, Dayan P (2014) The algorithmic anatomy of model-based evaluation. *Phil Trans R Soc B* 369:20130478.
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711.
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215.
- Dayan P, Niv Y (2008) Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol* 18:185–196.
- Decker JH, Lourenco FS, Doll BB, Hartley CA (2015) Experiential reward learning outweighs instruction prior to adulthood. *Cogn Affect Behav Neurosci* 15:310–320.
- Decker JH, Otto AR, Daw ND, Hartley CA (2016) From creatures of habit to goal-directed learners. *Psychol Sci* 27:848–858.
- Deserno L, Wilbertz T, Reiter A, Horstmann A, Neumann J, Villringer A, Heinze HJ, Schlagenhauf F (2015) Lateral prefrontal model-based signatures are reduced in healthy individuals with high trait impulsivity. *Transl Psychiatry* 5:e659–e659.
- Deserno L, Moran R, Lee Y, Michely J, Dayan P, Dolan R (2021) Dopamine enhances model-free credit assignment through boosting of retrospective model-based inference. *Biol Psychiatry* 89:594.
- Doll BB, Jacobs WJ, Sanfey AG, Frank MJ (2009) Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res* 1299:74–94.
- Doll BB, Hutchison KE, Frank MJ (2011) Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J Neurosci* 31:6188–6198.

- Doll BB, Bath KG, Daw ND, Frank MJ (2016) Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J Neurosci* 36:1211–1222.
- Eppinger B, Kray J (2011) To choose or to avoid: age differences in learning from positive and negative feedback. *J Cogn Neurosci* 23:41–52.
- Eppinger B, Walter M, Heekeren HR, Li SC (2013) Of goals and habits: age-related and individual differences in goal-directed decision-making. *Front Neurosci* 7:253.
- Ferreira D, Correia R, Nieto A, Machado A, Molina Y, Barroso J (2015) Cognitive decline before the age of 50 can be detected with sensitive cognitive measures. *Psychosoma* 27:216–222.
- Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci* 17:51–72.
- Frank MJ, Seeberger LC, O'reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 104:16311–16316.
- Fuhrmann D, Knoll LJ, Blakemore SJ (2015) Adolescence as a sensitive period of brain development. *Trends Cogn Sci* 19:558–566.
- Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, Paus T, Evans AC, Rapoport JL (1999) Brain development during childhood and adolescence: A longitudinal MRI study. *Nat Neurosci* 2:861–863.
- Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND (2016) Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife* 5:e11305.
- Gilmore JH, Shi F, Woolson SL, Knickmeyer RC, Short SJ, Lin W, Zhu H, Hamer RM, Styner M, Shen D (2012) Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cereb Cortex* 22:2478–2485.
- Gläscher J, Daw N, Dayan P, O'Doherty JP (2010) States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66:585–595.
- Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TF, Herman DH, Clasen LS, Toga AW, Rapoport JL, Thompson PM (2004) Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A* 101:8174–8179.
- Herzog N, Waltmann M, Deserno L, Janssen LK, Horstmann A (2023) Shared and differential neurocognitive mechanisms in obesity and binge eating disorder from adolescence to adulthood: An attempt to improve prediction of clinical outcome. Available at <https://doi.org/10.17605/OSF.IO/FYN6Q>.
- Hwang EJ, Dahlen JE, Mukundan M, Komiyama T (2017) History-based action selection bias in posterior parietal cortex. *Nat Commun* 8:1242.
- Jolles DD, Kleibeuken SW, Rombouts SAR, Crone EA (2011) Developmental differences in prefrontal activation during working memory maintenance and manipulation for different memory loads. *Dev Sci* 14:713–724.
- Knickmeyer RC, Gouttard S, Kang C, Evans D, Wilber K, Smith JK, Hamer RM, Lin W, Gerig G, Gilmore JH (2008) A structural MRI study of human brain development from birth to 2 years. *J Neurosci* 28:12176–12182.
- Kool W, Cushman FA, Gershman SJ (2016) When does model-based control pay off? *PLoS Comput Biol* 12:e1005090.
- Koolschijn PCMP, Schel MA, De Rooij M, Rombouts SAR, Crone EA (2011) A three-year longitudinal functional magnetic resonance imaging study of performance monitoring and test-retest reliability from childhood to early adulthood. *J Neurosci* 31:4204–4212.
- Kringelbach ML (2005) The human orbitofrontal cortex: linking reward to hedonic experience. *Nat Rev Neurosci* 6:691–702.
- Levy R, Goldman-Rakic PS (1999) Association of storage and processing functions in the dorsolateral prefrontal cortex of the nonhuman primate. *J Neurosci* 19:5149–5158.
- Louie K, De Martino B (2014) Neuroeconomics chapter 24: the neurobiology of context-dependent valuation and choice. In *Neuroeconomics: decision making and the brain* (Glimcher PW, Fehr E, eds) pp 455–476. Cambridge MA: Academic.
- Mills KL, Siegmund KD, Tamnes CK, Ferschmann L, Wierenga LM, Bos MGN, Luna B, Li C, Herting MM (2021) Inter-individual variability in structural brain development from late childhood to young adulthood. *Neuroimage* 242:118450.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Moran R, Keramati M, Dayan P, Dolan RJ (2019) Retrospective model-based inference guides model-free credit assignment. *Nat Commun* 10:750.
- Munoz F, Jensen G, Kennedy BC, Alkan Y, Terrace HS, Ferrera VP (2020) Learned representation of implied serial order in posterior parietal cortex. *Sci Rep* 10:9386.
- Nee DE, Brown JW, Askren MK, Berman MG, Demiralp E, Krawitz A, Jonides J (2013) A meta-analysis of executive components of working memory. *Cereb Cortex* 23:264–282.
- Niendam TA, Laird AR, Ray KL, Dean YM, Glahn DC, Carter CS (2012) Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn Affect Behav Neurosci* 12:241–268.
- Nussenbaum K, Hartley CA (2019) Reinforcement learning across development: what insights can we draw from a decade of research? *Dev Cogn Neurosci* 40:100733.
- Nussenbaum K, Scheuplein M, Phaneuf CV, Evans MD, Hartley CA (2020) Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra Psychol* 6:17213.
- Ott DVM, Ullsperger M, Jocham G, Neumann J, Klein TA (2011) Continuous theta-burst stimulation (cTBS) over the lateral prefrontal cortex alters reinforcement learning bias. *Neuroimage* 57:617–623.
- Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND (2013) Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A* 110:20941–20946.
- Palminteri S, Lebreton M (2021) Context-dependent outcome encoding in human reinforcement learning. *Curr Opin Behav Sci* 41:144–151.
- Palminteri S, Khamassi M, Joffily M, Coricelli G (2015) Contextual modulation of value signals in reward and punishment learning. *Nat Commun* 6:8096.
- Palminteri S, Kilford EJ, Coricelli G, Blakemore SJ (2016) The computational development of reinforcement learning during adolescence. *PLoS Comput Biol* 12:e1004953.
- Pauli R, Brazil I, Kohls G, Klein-Flügge MC, Rogers JC, Dikeos D, Dochnal R, Fairchild G, Fernández-Rivas A, Herpertz-Dahlmann B, Hervas A, Konrad K, Popma A, Stadler C, Freitag CM, De Brito SA, Lockwood PL (2022) Action initiation and punishment learning differ from childhood to adolescence while reward learning remains stable. *BioRxiv* 490578. <https://doi.org/10.1101/2022.05.05.490578>.
- Peters S, Van Duijvenvoorde ACK, Koolschijn PCMP, Crone EA (2016) Longitudinal development of frontoparietal activity during feedback learning: contributions of age, performance, working memory and cortical thickness. *Dev Cogn Neurosci* 19:211–222.
- Pischedda D, Palminteri S, Coricelli G (2020) The effect of counterfactual information on outcome value coding in medial prefrontal and cingulate cortex: from an absolute to a relative neural code. *J Neurosci* 40:3268–3277.
- Potter TCS, Bryce NV, Hartley CA (2017) Cognitive components underpinning the development of model-based learning. *Dev Cogn Neurosci* 25:272–280.
- Reitan RM (1955) The relation of the Trail Making Test to organic brain damage. *J Consult Psychol* 19:393–394.
- Reiter AMF, Deserno L, Wilbertz T, Heinze HJ, Schlagenhaut F (2016) Risk factors for addiction and their association with model-based behavioral control. *Front Behav Neurosci* 10:26.
- Roitman JD, Brannon EM, Platt ML (2007) Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol* 5:e208.
- Rosenbaum GM, Grassie HL, Hartley CA (2022) Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. *Elife* 11:e64620.
- Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB (2012) Modelling neural correlates of working memory: a coordinate-based meta-analysis. *Neuroimage* 60:830–846.
- Rushworth MFS, Noonan MAP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70:1054–1069.
- Schad DJ, Jünger E, Sebold M, Garbusow M, Bernhardt N, Javadi AH, Zimmermann US, Smolka MN, Heinz A, Rapp MA, Huys QJM (2014)

- Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Front Psychol* 5:1450.
- Schmidt K-H, Metzler P (1992) *Wortschatztest*. Weinheim: Beltz Verlag.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Silverman MH, Jedd K, Luciana M (2015) Neural networks involved in adolescent reward processing: an activation likelihood estimation meta-analysis of functional neuroimaging studies. *Neuroimage* 122:427–439.
- Singh-Manoux A, Kivimaki M, Glymour MM, Elbaz A, Berr C, Ebmeier KP, Ferrie JE, Dugravot A (2012) Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *BMJ* 344:d7622.
- Smid CR, Kool W, Hauser TU, Steinbeis N (2022) Computational and behavioral markers of model-based decision making in childhood. *Dev Sci* 26:e13295.
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. Cambridge, MA: MIT.
- Tamnes CK, Walhovd KB, Grydeland H, Holland D, Østby Y, Dale AM, Fjell AM (2013) Longitudinal working memory development is related to structural maturation of frontal and parietal cortices. *J Cogn Neurosci* 25:1611–1623.
- Tingley D, Yamamoto T, Hirose K, Keele L, Imai K (2014) Mediation: R package for causal mediation analysis. *J Stat Softw* 59:138.
- Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The neural basis of loss aversion in decision-making under risk. *Science* 315:515–518.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Vaghi MM, Moutoussis M, Váša F, Kievit RA, Hauser TU, Vértes PE, Shahr N, Romero-Garcia R, Kitzbichler MG, Bullmore ET, Dolan RJ (2020) Compulsivity is linked to reduced adolescent development of goal-directed control and frontostriatal functional connectivity. *Proc Natl Acad Sci USA* 117:25911–25922.
- Van Den Bos W, Cohen MX, Kahnt T, Crone EA (2012) Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cereb Cortex* 22:1247–1255.
- Van Der Schaaf ME, Warmerdam E, Crone EA, Cools R (2011) Distinct linear and non-linear trajectories of reward and punishment reversal learning during development: relevance for dopamine's role in adolescent decision making. *Dev Cogn Neurosci* 1:578–590.
- Van Leijenhorst L, Zanolie K, Van Meel CS, Westenberg PM, Rombouts SAR, Crone EA (2010) What motivates the adolescent? brain regions mediating reward sensitivity across adolescence. *Cereb Cortex* 20:61–69.
- Voon V, Derbyshire K, Ruck C, Irvine MA, Worbe Y, Enander J, Schreiber LRN, Gillan C, Fineberg NA, Sahakian BJ, Robbins TW, Harrison NA, Wood J, Daw ND, Dayan P, Grant JE, Bullmore ET (2015a) Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* 20:345–352.
- Voon V, Baek K, Enander J, Worbe Y, Morris LS, Harrison NA, Robbins TW, Ruck C, Daw N (2015b) Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Transl Psychiatry* 5:e670.
- Waltz JA, Frank MJ, Robinson BM, Gold JM (2007) Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry* 62:756–764.
- Wechsler D (1997) *WAIS-III. Wechsler Adult Intelligence Scale-third edition*. San Antonio, TX: The Psychological Association.
- Wierenga L, Langen M, Ambrosino S, van Dijk S, Oranje B, Durston S (2014) Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *Neuroimage* 96:67–72.
- Worbe Y, Palminteri S, Savulich G, Daw ND, Fernandez-Egea E, Robbins TW, Voon V (2016) Valence-dependent influence of serotonin depletion on model-based choice strategy. *Mol Psychiatry* 21:624–629.
- Zhong L, Zhang Y, Duan CA, Deng J, Pan J, Xu NL (2019) Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization. *Nat Neurosci* 22:963–973.
- Ziegler G, Hauser TU, Moutoussis M, Bullmore ET, Goodyer IM, Fonagy P, Jones PB, Lindenberger U, Dolan RJ (2019) Compulsivity and impulsivity traits linked to attenuated developmental frontostriatal myelination trajectories. *Nat Neurosci* 22:992–999.