Learning in anticipation of reward and punishment: Perspectives across the human lifespan

Matthew J. Betts[1,2,13#], Anni Richter[3,#], Lieke de Boer[4], Jana Tegelbeckers[5,6], Valentina Perosa[2,10], Valentin Baumann[6], Rumana Chowdhury[7], Ray J. Dolan[8,9], Constanze Seidenbecher[3,13], Björn H. Schott[3,10,11,13,14], Emrah Düzel[1,2,12,13], Marc Guitart-Masip[4,9,*], Kerstin Krauel[6,13,*]

[1]German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany
[2]Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
[3]Department of Neurochemistry and Molecular Biology, Department of Behavioral Neurology, Leibniz Institute for Neurobiology, Magdeburg, Germany
[4]Aging Research Centre, Karolinska Institute, Stockholm, Sweden
[5] Feinberg School of Medicine, Northwestern University, Chicago, United States
[6]Department of Child and Adolescent Psychiatry and Psychotherapy, Otto von Guericke University, Magdeburg, Germany
[7]Leeds Teaching Hospital NHS Trust, Leeds, UK
[8]The Wellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London, WC1N 3BG, UK
[9]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London
[10]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany
[11]Department of Neurology, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
[12]Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK
[13]Center for Behavioral Brain Sciences, University of Magdeburg, Magdeburg, Germany
[14]German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany

[#,] *Denotes equal contribution

Corresponding authors: Matthew J. Betts and Anni Richter

Key words: action, reward, Pavlovian bias, instrumental learning, lifespan

Abstract

Learning to act to receive reward and to withhold to avoid punishment has been found to be easier than learning the opposite contingencies in young adults. To what extent this type of behavioral adaptation might develop during childhood and adolescence and differ during aging remains unclear. We therefore tested 247 healthy individuals across the human lifespan (7-80 years) with an orthogonalized valenced go/no-go learning task. Computational modeling revealed that peak performance in young adults was attributable to greater sensitivity to both reward and punishment. However, in children and adolescents, we observed an increased bias towards action but not reward sensitivity. In contrast, reduced learning in midlife and older adults was accompanied by decreased reward sensitivity and especially punishment sensitivity along with an age-related increase in the Pavlovian bias. These findings reveal distinct motivation-dependent learning capabilities across the human lifespan, which cannot be probed using conventional *go*/reward *no-go*/punishment style paradigms that have important implications in life-long education.

Introduction

Flexible instrumental learning is required to adapt behavior towards maximizing reward and minimizing punishment. A number of studies have shown that its success critically depends on whether reward or avoidance of punishment is paired with action or inhibition of action, respectively. When using a go/no-go task that independently dissociates, i.e. orthogonalizes, action and valence, young adults demonstrate striking asymmetry in instrumental learning – signals that predict reward are prepotently associated with behavioral activation, whereas signals that predict punishment are intrinsically coupled to behavioral inhibition (Guitart-Masip *et al.*, 2012*b*; Cavanagh *et al.*, 2013, Chowdhury *et al.*, 2013*b*; Guitart-Masip *et al.*, 2014; Richter *et al.*, 2014; Perosa *et al.*, 2020). Considering previously reported differences in neural correlates of reward and punishment processing in children and adolescents (van Duijvenvoorde *et al.*, 2008) as well as in elderly adults (Samanez-Larkin *et al.*, 2007; Schott *et al.*, 2007; de Boer *et al.*, 2017), it is crucial to consider both developmental and aging effects on this flexible but also biased behavior for the formulation of comprehensive instrumental learning theories and for lifelong education.

Valence may interact with action and bias performance at two different time points in our task. Anticipated value may bias action selection such that positive valence promotes action and negative valence promotes inhibition. This is a Pavlovian mechanism and the observed effect in our task is termed a Pavlovian bias (Huys *et al.*, 2011, Guitart-Masip *et al.*, 2012*b*; Geurts *et al.*, 2013). Valence can also interact with action during outcome processing resulting in boosted or reduced learning after a rewarded actions or punished inaction, which would be considered an instrumental mechanism (Swart *et al.*, 2017) referred to here as an instrumental learning bias. Both mechanisms have found support in previous studies (Cavanagh *et al.*, 2013, Chowdhury *et al.*, 2013*b*; Guitart-Masip *et al.*, 2013; Richter *et al.*, 2014; Swart *et al.*, 2017; de Boer *et al.*, 2019; Perosa *et al.*, 2020). Learning success in this go/no-go task requires flexibility, inhibition, and the ability to use feedback and to detect reward contingencies. All of these abilities may critically rely on prefrontal cortex (PFC)-dependent executive functions, which develop during adolescence but also decline in older age (Kray *et al.*, 2004; Zelazo *et al.*, 2004). Whilst most of the studies to date using the aforementioned task have investigated young adults, little attention has focused on how the processes underlying instrumental learning and

potential conflict with Pavlovian mechanisms may change across the lifespan. Previous work has reported age-related differences in Pavlovian and instrumental learning biases (Chowdhury *et al.*, 2013*b*; de Boer *et al.*, 2019; Perosa *et al.*, 2020), but it has not been investigated yet whether differences already occur during midlife. Considering the other end of the lifespan, it is further unclear to what extent age-related differences in the Pavlovian bias reflect effects of maturation versus senescence. The inclusion of children and adolescents can help to resolve this open question.

It has been widely stated that adolescents are highly sensitive to reward, which may contribute to increased risky behavior during this developmental period (Casey *et al.*, 2008). However, it has also been suggested that this reward sensitivity may be adaptive by promoting learning and exploration — critical for transition into adulthood (Spear, 2000; Casey, 2015). A recent study demonstrated that adolescents learn to preferentially seek rewards rather than to avoid punishments, whereas young adults learn both behaviors equally well (Davidow *et al.*, 2016). However, previous studies have not dissociated reward sensitivity from action learning, and it remains thus unclear if this interpretation may be confounded by action requirements or to what extent changes in reward sensitivity may influence the strength of coupling between action and valence. This is particularly relevant in light of differential functional and anatomical development of limbic regions, such as the striatum and cognitive control regions during adolescence (Blakemore and Robbins, 2012; Shulman *et al.*, 2016). Such asymmetrical development may translate into differential Pavlovian and instrumental strategies used by children and adolescents compared to those employed in adulthood and may have important implications in neurodevelopmental disorders such as attention deficit hyperactivity disorder (Kuo and Liu, 2019).

The human brain also undergoes substantial change during normal aging, which has been associated with numerous cognitive changes (Bäckman *et al.*, 2006; Lindenberger, 2014). However, it is not known how age-related differences in Pavlovian and instrumental control mechanisms may impact behavioral inflexibility in older adults or whether such age-related changes are already evident in midlife adulthood – an age range mostly neglected in aging studies. Previous work has shown that administration of the dopamine precursor L-DOPA enhances the Pavlovian influence of potential reward (Rutledge *et al.*, 2015) and may restore reward prediction errors in old age (Chowdhury *et al.*, 2013*a*).

Coupled with the known age-related decline in the integrity of the dopaminergic system (Karrer *et al.*, 2017), a loss of functional dopamine may lead to a decrease in Pavlovian control in older age. Alternatively, previous studies have shown that the PFC is involved in overcoming the Pavlovian bias in young adults (Guitart-Masip *et al.*, 2012*b*; Cavanagh *et al.*, 2013). Thus, decreased functionality of the PFC as a result of normal aging could also lead to increased Pavlovian biases in older adults.

The objective of this study was to explore how acquisition of optimal/adaptive behavioral choices is differentially altered across the lifespan in 247 healthy participants (age 7-80 years) using an established go/no-go task that orthogonalizes action and valence. In addition, we used computational modeling to investigate different parameterizations on each subject's behavior to assess how Pavlovian and instrumental biases influence flexible learning across the lifespan and vary with age.

Materials and Methods

Participants

247 individuals between 7 to 80 years of age participated in the current study. Overall, the sample consisted of 111 females (44,9%) and 136 males (55,1%). It was ensured either by a standardized telephone interview or personal clinical interview that none of the participants were affected by a present or past neurological or psychiatric illness, alcohol or drug abuse, or were using centrally acting medication. Cognitive abilities were explicitly assessed in children, adolescents and older adults (> 60 years) to ascertain they had intact global cognitive performance (for details see SI Materials and Methods). Adult participants were only included if they had finished compulsory education (minimum 12 years). All participants received detailed oral and written information about the study and gave written consent. For minors, informed consent from children and adolescents as well as their parents was required for participation. The study was approved by the local ethics committee of the University of Magdeburg, Faculty of Medicine, and followed the ethical standards of the Declaration of Helsinki.

Task and Procedure

All participants performed a previously established valenced go/no-go probabilistic learning task (Guitart-Masip *et al.*, 2012*b*). Participants had to learn through trial and error, which of four fractal cues, preceding an easy visual target detection task, indicated the need (1) to respond to obtain a monetary reward (*go to win*), (2) to respond to avoid a monetary loss (*go to avoid losing*), (3) to withhold a response to obtain a monetary reward (*no-go to win*) and (4) to withhold a response to avoid a monetary loss (*no-go to avoid losing*) (see Figure 1). After display of the fractal cue (1000 ms), participants were presented with the target detection task (1500 ms). During the visual target detection task, participants were presented with a circle either on the right or left side of the screen and had to decide whether they should indicate (*go*) the target side or refrain from pressing a button (*no-go*). For the *go* conditions, they had to press a button indicating the side of the target within 2000 ms. Following the circle, participants obtained one of the following feedbacks (1000 ms): a green up-pointing arrow indicating a win (30 cent in children and adolescents or 50 cent in the adult groups), a red down-pointing arrow indicating a loss of 30/50 cents, or a yellow horizontal bar representing neither win nor loss. Feedback was probabilistic, thus, in the win conditions 80% of correct choices and 20% of incorrect choices were rewarded. In the lose conditions, 80% of correct choices and 20% of incorrect choices successfully avoided loss. Participants were informed and instructed about the probabilistic nature of the task beforehand.

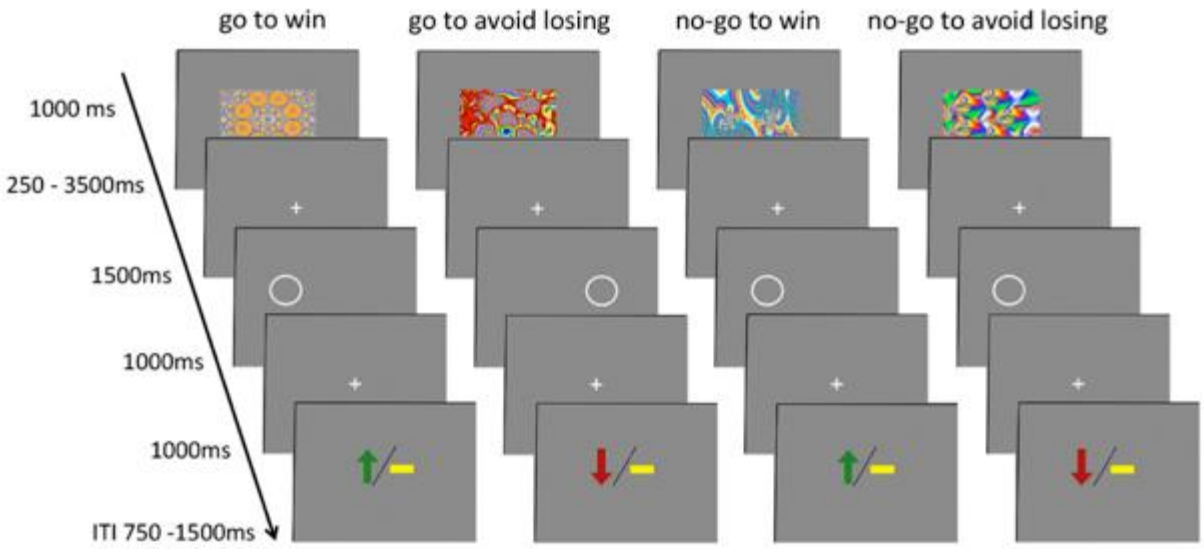

Figure 1: Probabilistic monetary go/no-go task. Participants had to learn over the course of 60 trials per condition, which fractal image was associated with responding or withholding a response to achieve a successful outcome (win or avoid losing). Responses

indicated whether the circle presented after the fractal was located on the right or left side. Correct feedback was only provided in 80% of the trials. Abbreviations: ITI – inter-trial interval. Figure adopted from Richter et al., 2014.

The task consisted of 240 trials (60 trials for each of the four conditions, presented in a randomized fashion in four runs) and lasted approximately 35 minutes. Before the task, participants were asked to complete 10 practice trials in which only the target detection circles were presented to familiarize themselves with the appropriate buttons on the computer keyboard and to obtain an overall feel for the speed of the task without exposure to any of the fractal cues used in the main task. The possible win/loss per trial for adults was 0,50 €. Children and adolescents received a slightly lower reward (0.30 Euro) to limit the maximum cumulative amount of money they could gain on completion of the task (children/adolescents: 21.60 Euro vs. adults 36 Euro). Moreover, reimbursement and the final reward were given in the form of gift vouchers (5 €) for a local shopping center in line with our previous studies and recommendations from the local Ethics committee on how to provide age-appropriate level reimbursement for school children. Adults received the exact amount they won or maximum of 25 Euro on completion of the task whereas for children and adolescents, earnings were rounded to 5 or max 10 Euro gain. Stimuli were presented and responses recorded using the Cogent 2000 toolbox running on MATLAB (Version 2009b; Mathworks).

Behavioral data analysis

To inspect and visualize trajectories of performance measures and modeling parameters across age, smooth curves were fitted via locally estimated scatterplot smoothing (LOESS) using the loess function in R (R Core Team, 2020) with a smoothing span of $\alpha = .75$ and a polynomial degree of $k = 2$. Visualizations of the LOESS curves were created via the ggplot package (Wickham, 2016).

For the behavioral data analysis, SPSS Advanced Statistics v26 (IBM Corporation, Armonk, NY, USA) was used. Initially, we assessed whether a Pavlovian bias could be detected in the sample. Thus, mean accuracy rates (percentage of correct responses, %) were analyzed in a two-factorial ANCOVA for repeated measures with the factors *action* (*go* vs. *no-go*) and *valence* (*win* vs. *avoid losing*). Since gender distributions were not uniform across age, gender was included as a covariate of no interest.

To analyze differential age effects on overall performance and acquisition in the four conditions *go to win (GoW), go to avoid losing (GoL), no-go to win (NoGoW), no-go to avoid losing (NoGoL)*, we performed polynomial regression analyses employing fits with linear and quadratic models. In detail, we calculated whether mean accuracy and learning gain (mean accuracy over the last 30 trials – mean accuracy over the first 30 trials) in all four conditions could be predicted by age. To account for multiple testing (n = 8, a Holm-Bonferroni correction was applied to p (change)-values.

Reinforcement learning models

We fitted choice behavior to a set of six nested reinforcement learning (RL) models incorporating different RL hypothesis. The base model was a Q-learning algorithm (Sutton and Barto, 1998) that used a Rescorla-Wagner update rule to independently track the action value of each choice given each fractal image ($Q_t(go)$ and $Q_t(nogo)$), with a learning rate (ε) as a free parameter. In the model, the probability of choosing one action on trial t was a sigmoid function of the difference between the action values scaled by a slope parameter that was parameterized as sensitivity to reward. This basic model was initially augmented with an irreducible action noise parameter also known as a lapse rate (ξ) (Talmi *et al.*, 2008) and then further expanded by adding a static bias parameter to the value of the go action (b). Note that the parameterization of the irreducible noise parameter implies that higher values are associated with lower irreducible noise. Hence this parameter can be interpreted as a lapse rate. The model was then augmented by adding a fixed Pavlovian value of 1 to the value of the go action as soon as the first reward was encountered for win cues, and a fixed Pavlovian value of -1 to the value of the go action as soon as the first punishment was encountered for loss cues. This fixed Pavlovian value was weighted by a further free parameter (Pavlovian parameter) into the value of the go action (π). Note that this definition of the Pavlovian value is different from the definition in previous studies that have used this task (Guitart-Masip *et al.*, 2012*b*; Cavanagh *et al.*, 2013; de Boer *et al.*, 2019), as model comparison demonstrated it a better fit than a variable Pavlovian value updated on a trial-by-trial basis (see Table 1). The state (action-independent) values for each fractal image were updated on every trial using a Rescorla-Wagner update rule with the same learning rate as the update of the action values. Finally, the model including the static action bias and the Pavlovian bias were augmented by including different sensitivities for reward and punishment. Full equations and a description of all considered models are provided in the Supplemental Information.

Model fitting procedure and comparison

As in previous reports (Huys *et al.*, 2011, Guitart-Masip *et al.*, 2012*b*) we used a hierarchical Type II Bayesian (or random effects) procedure using maximum likelihood to fit simple parameterized distributions for higher-level statistics of the parameters. Since the values of parameters for each subject are 'hidden', this employs the Expectation-Maximization (EM) procedure. For each iteration, the posterior distribution over the group for each parameter is used to specify the prior over the individual parameter fits on the next iteration. All six computational models were fit to the data using a single distribution for all participants. This fitting procedure was, therefore, blind to the existence of different groups with putatively different parameter values. During the fitting procedure, all parameters except the action bias were suitably transformed to enforce constraints (exponential transform for sensitivity to reward and punishment and Pavlovian parameter and sigmoid transforms for learning rate and irreducible noise). These constraints are enforced to ensure stability of the model fitting as recommended by Daw 2010. The learning rate and the irreducible noise parameters are fractional step sizes and naturally range between 0 and 1; numbers below 0 are meaningless, and numbers above 1 render the estimation of the model unstable. Similarly, negative values of the sensitivity to reward, the sensitivity to punishment, and the Pavlovian parameter are logically implausible. Six modeling parameters were extracted for each individual, namely *sensitivity to reward*, *sensitivity to punishment, Pavlovian bias, action bias, learning rate* and *irreducible noise.*

Models were compared using the integrated Bayesian Information Criterion (iBIC) as previously described (Huys *et al.*, 2011, Guitart-Masip *et al.*, 2012*b*). Small iBIC values indicate a model that fits the data better after penalizing for the number of data points associated with each parameter. Comparing iBIC values is akin to a likelihood ratio test (Kass and Raftery, 1995). Note that the iBIC penalizes those versions of the model fit that use four distributions for each parameter. Finally, we assessed age-related effects on all modeling parameters using polynomial regression analyses employing fits with linear and quadratic models. Holm Bonferroni-correction ($p < 0.05$ for 6 tests) was used to correct for effects of multiple comparisons.

Data and Code availability statements

The data and code used to support the findings of this study are available from the corresponding authors upon reasonable request.

Results

Pavlovian bias

Across the entire sample, participants showed superior performance in the *go* conditions (main effect *action*: $F_{1,244} = 26.618$, $p < .001$). However, the significant interaction of *action* x *valence* ($F_{1,244} = 14.009$, $p < .001$) revealed the well-described Pavlovian bias, namely that performance in the *go* conditions was higher when a reward was expected (*go to win < go to avoid losing,* $t_{246} = 10.453$, $p < .001$), whereas the prospect to avoid a loss led to better performance in the *no-go* conditions (*no-go to win < no-go to avoid losing,* $t_{246} = -7.817$, $p < .001$). No significant interactions with gender were observed (all $p \geq 0.5$). Figure 2 shows the Pavlovian Bias and its trajectory across the lifespan.

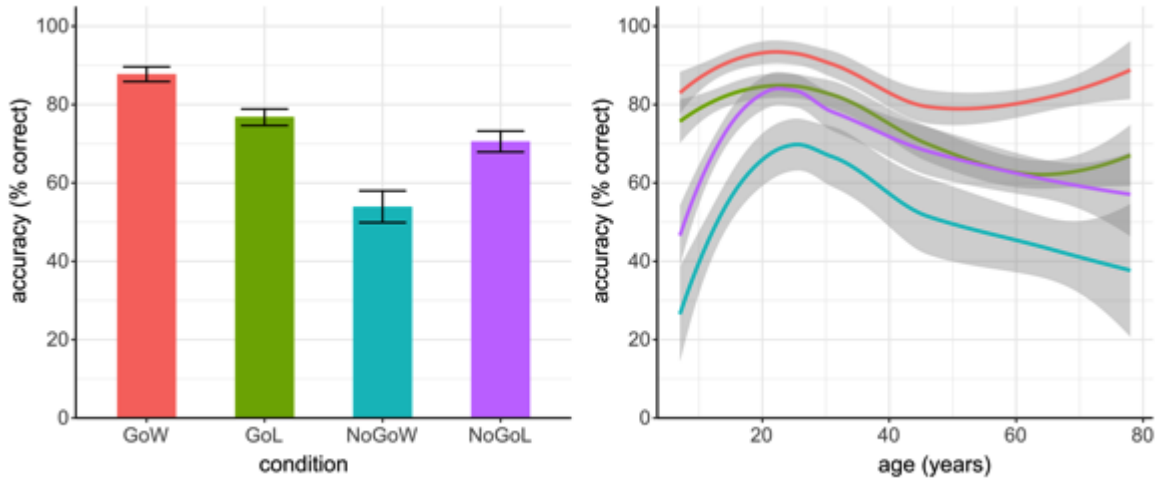

Figure 2: Left: Mean overall accuracy (%) with 95% confidence intervals across conditions indicating a Pavlovian bias during learning (*GoW > GoL and NoGoW < NoGoL,* ***p<0.001)*;* Right: Trajectories of overall accuracy (%) for all conditions plotted as a function of age. Smooth curves fitted using locally estimated scatterplot smoothing (LOESS) are shown with 95% confidence interval. Abbreviations: GoW, *go to win*; GoL, *go to avoid losing*; NoGoW, *no go to win*; NoGoL, *no go to avoid losing*.

Age related effects on performance accuracy

In a first step, we looked at the four conditions separately (see Figure 3 and Table 1 upper parts): performance in the *go* conditions decreased linearly with age, both for *go to win*

(F(1,245) = 7.716, p = .006) and *go to avoid losing* (F(1,245) = 55.018, p < .001). In contrast, performance in the *no-go* conditions showed a non-linear (quadratic) relationship with age, that is, rising steeply during childhood and adolescence, peaking in young adulthood and then showing a continuous age-related decrease in midlife and older adults (*no go to win*: F(1,244) = 12.149, p < .001; *no go to avoid losing:* F(1,244) = 17.96, p < .001).

Considering increments in performance from the first to the second half of trials (learning gain, see bottom rows of Figure 3 and Table 1), we observed no age effects for the *go* conditions. In the *no-go to win condition*, learning gain was higher in children and adolescents as well as young adults and decreased linearly with age (F(1,245) = 8.834, p = .003). Learning gain in the *no-go to avoid losing* condition, however, showed a moderate non-linear/quadratic association with age (F(1,244) = 7.720, p = .001), showing a slight age-related increase up to midlife and subsequent age-related decrease in older age.

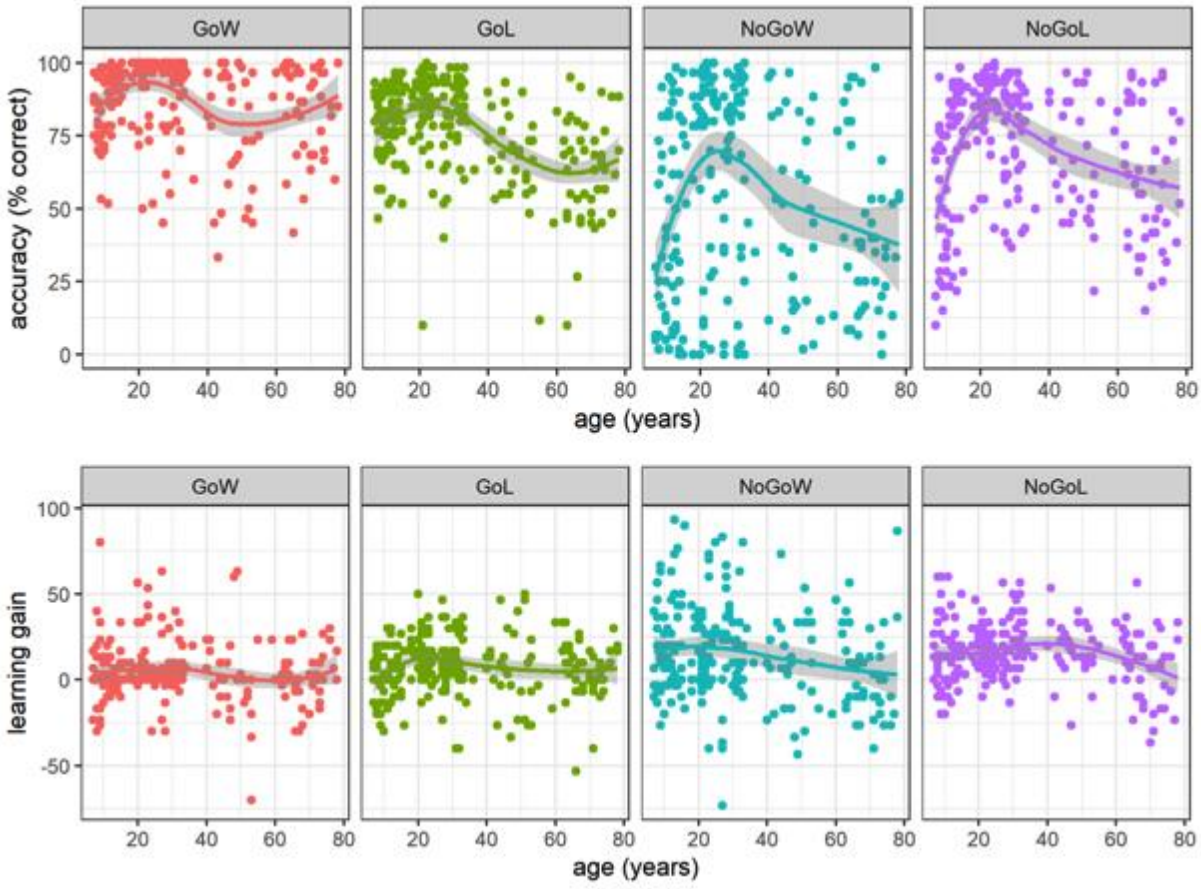

Figure 3: Mean performance of each participant across all conditions. Smooth curves using LOESS are shown with 95% confidence intervals: Upper row: Mean accuracy (%). Bottom row: Learning gain: mean accuracy (%) last 30 trials – mean accuracy (%) first 30

trials. Abbreviations: GoW, *go to win*; GoL, *go to avoid losing*; NoGoW, *no go to win*; NoGoL, *no go to avoid losing*.

For completeness, age effects on reaction times for *go to win* and *go to avoid losing* responses as well as inverse efficiency scores (i.e. RT/performance accuracy) are reported in the supplementary information (SI Figure 1). Their interpretation warrants caution as participants were explicitly instructed to respond accurately, while speed was not emphasized. Reaction times showed a highly significant quadratic association with age (GoW: $F(1,244) = 58.289$, $p < .001$, GoL: $F(1,244) = 59.699$, $p < .001$) as was the case for the inverse efficiency for *go to win* ($F(1,244) = 23.161$, $p < .001$), indicating that young adults showed faster responses (steep increase during childhood and adolescence) and a lower speed-accuracy trade off. For inverse efficiency scores in the *go to avoid losing* condition, a significant linear association with age was observed ($F(1,245) = 31.567$, $p < .001$)

Table 1: Association of accuracy measures with age

		Linear		Quadratic	
		R^2 (change)	p (change)	R^2 (change)	p (change)
Accuracy	GoW	**.031**	**.006**	.001	.540
	GoL	**.183**	**<.001**	.017	.023
	NoGoW	.007	.189	**.084**	**<.001**
	NoGoL	.020	.025	**.108**	**<.001**
Learning gain	GoW	.009	.138	.001	.563
	GoL	.001	.629	.019	.029
	NoGoW	**.035**	**.003**	.003	.394
	NoGoL	.020	.028	**.040**	**.001**

Results that survived Holm-Bonferroni correction are shown in bold print.

Parameterizing learning and biases using computational modeling

To identify instrumental and Pavlovian components of the observed asymmetry during learning, six nested reinforcement learning (RL) models were fitted to the behavioral data (see SI Materials and Methods), using the expectation maximization approach as previously described (Huys *et al.*, 2011, Guitart-Masip *et al.*, 2012*b*). All six computational

models were fit to the data using a single distribution for all participants. This fitting procedure was, therefore, blind to the existence of different groups with putatively different parameter values. Our computational modeling approach demonstrated that the marked asymmetry in learning (i.e. superior performance in *go to win* and *no-go to avoid losing* compared to *go to avoid losing* and *no-go to win*) could be attributed to an interaction between instrumental and Pavlovian control mechanisms (Guitart-Masip *et al.*, 2012*b*). The best account of the data was provided by the model including a static *action bias*, *Pavlovian bias*, *reward* and *punishment sensitivity, learning rate* and an *irreducible noise* parameter (see Table 2 and SI Figure 2) consistent with previous studies using this task (Guitart-Masip *et al.*, 2012*b*; Cavanagh *et al.*, 2013; Guitart-Masip *et al.*, 2014; Perosa *et al.*, 2020).

Table 2: Integrated Bayesian Information Criterion (iBIC) for all tested models

Model no.	Model parameters	No. of parameters	Likelihood	Pseudo-R^2	iBIC
1	ε, ρ	2	-28166	0.32	56,376
2	ε, ρ, ξ	3	-27977	0.32	56,019
3	ε, ρ, ξ, b	4	-25653	0.38	51,394
4	ε, ρ_{win}, ρ_{lose}, ξ, b	5	-25064	0.39	50,239
5	ε, ρ_{win}, ρ_{lose}, ξ, b, π_{fluct}	6	-24680	0.40	49,493
6	**ε, ρ_{win}, ρ_{lose}, ξ, b, $\pi_{constant}$**	**6**	**-24519**	**0.41**	**49,170**

The winning model statistics are highlighted in bold font: $\sum$, learning rate; **ρ**win, weighting of reward on win trials; **ρ**lose, weighting of punishments on lose trials; **b**, go bias; **π**, Pavlovian bias; **ξ**, irreducible noise. iBIC, integrated Bayesian information criterion. Smaller values indicate a better model fit.

Age-related effects on modeling parameters

Parameters indicative of instrumental learning such as *reward sensitivity* ($F(1,244) = 9.773$, $p < .001$) and *punishment sensitivity* ($F(1,244) = 19.084$, $p < .001$) both followed an inverted U-shaped distribution across the lifespan (see upper parts of Figure 4 and Table 3). An age-related increase in *reward* and *punishment sensitivity* was observed during childhood and adolescence, peaking during young adulthood followed by an age-related decrease in midlife and older adults. In contrast, no significant association between *learning rate* and age across the lifespan was observed.

Moreover, we found a nonlinear/quadratic association between age and *action bias* ($F(1,244) = 9.389$, $p < .001$; see lower parts of Figure 4 and Table 3) showing that

associative learning in children and adolescents was driven by an *action bias* that attenuated in young and midlife adults and recurred in older age. The modeling approach also revealed a significant linear association between the *Pavlovian bias* and age ($F(1,245) = 6.270$, $p < .013$) indicating that associative learning is more influenced by the *Pavlovian bias* with increasing age. For the *irreducible noise* parameter, we observed a significant linear association with age ($F(1,245) = 8.854$, $p < .003$) whereby values were highest in young adults and lowest in older adults demonstrating younger adults' performance was more tightly captured by the winning model.

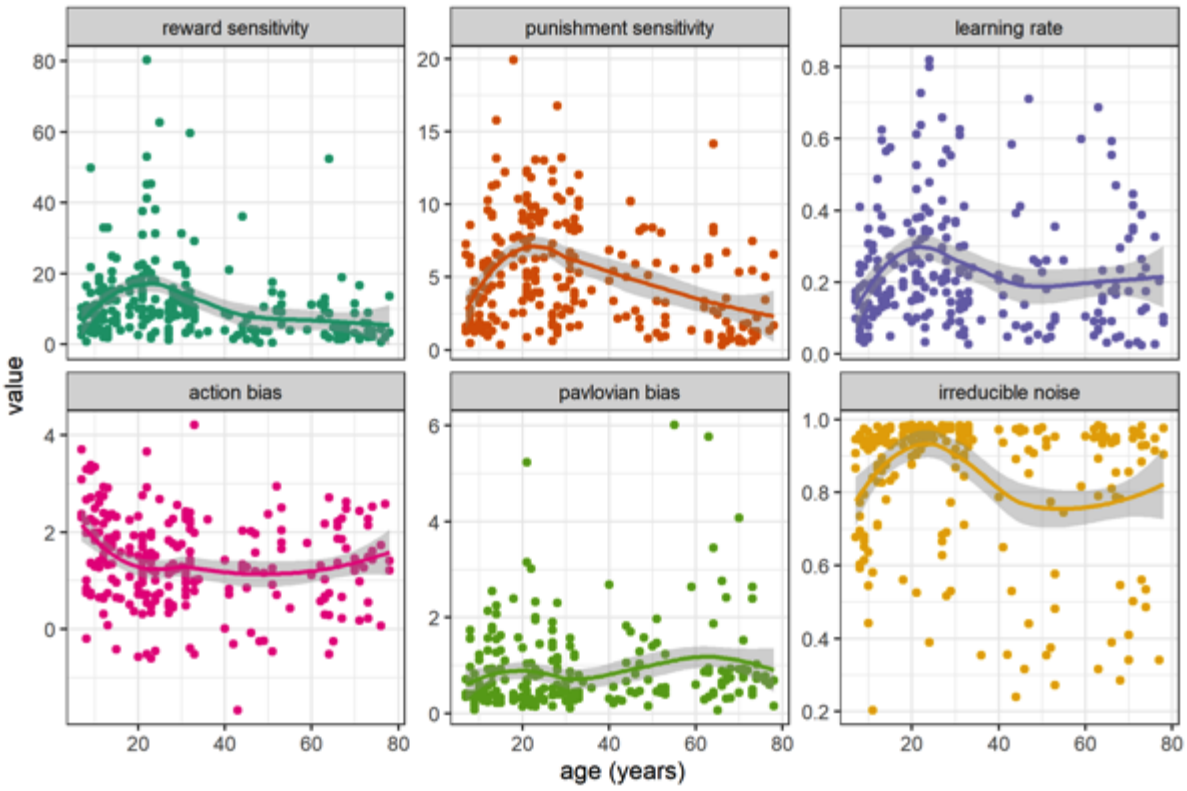

Figure 4: Modeled behavioral performance across the lifespan. Modeling parameters derived from the winning model are plotted as smooth curves using LOESS and are shown with 95% confidence intervals.

Table 3: Association of modeling parameters with age

	Linear		Quadratic	
	R^2 (change)	p (change)	R^2 (change)	p (change)

Modeling parameters	Reward sensitivity	**.048**	**.001**	**.026**	**.009**
	Punishment sensitivity	**.060**	**<.001**	**.075**	**<.001**
	Learning rate	.003	.419	.018	.036
	Action bias	.019	.029	**.052**	**<.001**
	Pavlovian bias	**.025**	**.013**	.000	.967
	Irreducible noise	**.035**	**.003**	.011	.088

Results that survived Holm-Bonferroni correction are shown in bold print.

Discussion

In this study we have demonstrated, that individual performance in a valenced go/no-go task across the lifespan (7-80 years) is influenced by *Pavlovian biases*. Furthermore, the ability to successfully orthogonalize action and valence was characterized by an inverted U-shape distribution with peak performance observed in young adults. Computational modeling revealed that superior performance in younger adults compared to all other age groups was attributable to a greater sensitivity to outcomes (both to reward and punishment) coupled with a relatively low *action bias*. In contrast, lower performance in children and adolescents was attributable to an increased bias towards action and reduced *reward* and *punishment sensitivity* compared to young adults. In midlife and older adults, an age-related decline in performance was attributable to a decrease in *reward sensitivity* and especially *punishment sensitivity* and age-related increase in the *Pavlovian bias*. Taken together, our results demonstrate that age-related decline in motivated performance does not mirror the pattern observed in children and adolescents, but instead reflects qualitatively distinct alterations. Our study thus reveals novel age-related discrepancies that could not be probed using classical *go*/reward *no-go*/punishment style paradigms.

This study set out to investigate the influence of Pavlovian biases on instrumental learning responses coupled to action and valence using a go/no-go task. Our results revealed that all participants exhibited an influence of Pavlovian control whereby they were better at initiating an action to gain a reward (*go to win*) compared to punishment (*go to avoid*

losing) but also withdrawing an action to avoid punishment (*no-go to avoid losing*) compared to gaining a reward (*no-go to win*). The striking asymmetry in performance across conditions observed here is consistent with previous studies using the same task (Guitart-Masip *et al.*, 2012*b*; Cavanagh *et al.*, 2013, Chowdhury *et al.*, 2013*b*; Guitart-Masip *et al.*, 2014; Richter *et al.*, 2014; de Boer *et al.*, 2019; Perosa *et al.*, 2020). Furthermore, computational modeling in young adults has previously shown this pattern of behavior can be captured by a model incorporating a Pavlovian bias, where the strength of this bias is related to impaired learning of the conflicting conditions: *no-go to win* and *go to avoid losing* (Guitart-Masip *et al.*, 2012*b*). Here we extend this work by demonstrating that the same model can effectively capture learning behavior across the human lifespan. Interestingly, we observed a significant age-related increase in the *Pavlovian bias* in older age providing evidence for age-differential Pavlovian control. However we also show that performance is influenced by age-related differences in the ability of the instrumental system to learn the appropriate choice (go or no-go) for each fractal image, as observed by significant age-related decreases in *reward* and *punishment sensitivity* but not *learning rate* in older age.

While children and adolescents demonstrated Pavlovian responding consistent with all other age groups, we observed an underlying preference for action, that is for *go* responses regardless of valence, which is at odds with the prevalent view that there is an overall increase in reward sensitivity during adolescence (for reviews see Galvan, 2010; Walker *et al.*, 2017). Instead, our results rather suggest that this is only true for rewards coupled to action. In fact, our computational modeling analysis revealed that children and adolescents showed lower rather than higher *reward sensitivity* when compared to younger adults. As this is the first study to our knowledge to dissociate or orthogonalize action and valence during motivated learning in children and adolescents, our results raise the possibility that increased reward sensitivity reported in previous studies (Galvan *et al.*, 2006; Van Leijenhorst *et al.*, 2010; Somerville *et al.*, 2011; Palminteri *et al.*, 2015; Peeters *et al.*, 2017) might actually be attributable to a preference for action responses, which were coupled to positive outcomes. When considering these implications, it becomes evident that, in the light of our findings, it is important to control for action tendencies when investigating reward learning in children and adolescents. Previous work in young adults using the same *go*/*no-go* task has shown that activity in inferior frontal gyrus (IFG), a region known to be involved in action inhibition (Aron *et al.*,

2014), is associated with *no-go* learning and successful instrumental control (Guitart-Masip *et al.*, 2012*b*). Considering the previously described protracted development of prefrontal projections to subcortical regions (Ziegler *et al.*, 2017), we tentatively suggest that lower PFC-dependent top-down regulation of the striatal reward system may contribute to the increased action bias in children and adolescents.

In older adults, we found poorer overall performance in all task conditions, but also a considerably reduced ability to learn the stimulus-response associations across trials when compared to young adults. A study by Schott *et al.*, (2007) suggested that, compared to young adults, older adults exhibited profoundly reduced mesolimbic activation during reward anticipation, although did activate the ventral striatum during reward feedback. Similarly, Chowdhury *et al.*, (2013*a*) demonstrated that older adults do not show a representation of expected value in the ventral striatum when performing a probabilistic reward learning under basal conditions and that an expected value representation was only observed after boosting the dopaminergic system with L-DOPA. Furthermore, older adults performing a probabilistic reward learning task show an attenuation of value anticipation in the ventromedial prefrontal cortex (vmPFC) that predicts performance in the probabilistic learning task (de Boer *et al.*, 2017). These findings suggest that whilst general reward processing may be intact in older adults, they are impaired in learning the predictive value of probabilistic reward cues. In fact, previous studies have shown that reduced learning in older adults is associated with deficits in the integration and updating of reward information when rewards are uncertain and delivered from probabilistic outcomes (Eppinger *et al.*, 2008; Hämmerer *et al.*, 2011). Additionally, Samanez-Larkin *et al.*, (2007) showed a particularly strong age-related impairment of ventral striatal loss anticipation in older adults. Taken together, these findings are compatible with our current observation of decreased instrumental learning in older adults, which was attributable to an age-related decrease in *reward* and particularly *punishment sensitivity* in the learning model.

A striking finding in our data is that the impairment of the instrumental learning in older adults was especially manifest as decreased performance in the *go* conditions. The dopamine system is involved in generating active motivated behavior (Niv *et al.*, 2007; Salamone and Correa, 2012) and instrumental learning through reward-prediction errors (Schultz, 2010). Dopamine depletion leads to decreased motor activity and/or reduced

motivation to seek rewards (Palmiter, 2008; Salamone and Correa, 2012). Similarly, previous findings using a variant of the same *go*/*no-go* task have shown that L-DOPA administration invigorated instrumental responding regardless of valence (Guitart-Masip *et al.*, 2012*a*). Therefore, an age-related decline in dopaminergic function as previously described (Bäckman *et al.*, 2006; Karrer *et al.*, 2017), could modulate motivation or vigor of actions independently of valence and may explain the overall decrease in *go* performance observed in older compared to younger adults.

Determining the impact of an aging dopaminergic system on performance in the valenced go/no-go task is not straightforward. Most previous studies support the notion that dopamine facilitates the action by valence interaction during learning (see however Guitart-Masip *et al.,* 2014). A study investigating a genetic variant linked to dopamine D2 receptor expression also highlights a modulatory role for genetic variability within the dopaminergic system in individual learning differences of action-valence interactions (Richter *et al.*, 2014). Another study has shown that boosting dopamine with methylphenidate increases the action by valence interaction in participants with high working memory capacity, a proxy for higher dopamine synthesis capacity (Swart *et al.*, 2017). Finally, a recent PET imaging study has shown that the strength of the action by valence interaction scales with the availability of dopamine D1 receptors in the dorsal striatum independent of age (de Boer *et al.*, 2019). Based on this evidence, one would have predicted a decreased Pavlovian bias in older age. However, learning the correct contingencies of the orthogonalized go/no-go task may critically rely on high-level cognitive functions, thus lifespan differences reported here may also relate to interindividual differences in working memory and long-term memory. These cognitive functions may also be compromised as a result of age-related decline in grey and white matter integrity (Draganski *et al.*, 2011; Samanez-Larkin *et al.*, 2012, Chowdhury *et al.*, 2013*b*; Callaghan *et al.*, 2014; Acosta-Cabronero *et al.*, 2016; Steiger *et al.*, 2016; van de Vijver *et al.*, 2016; Perosa *et al.*, 2020) which may influence the ability of the instrumental system to learn the task contingencies as indexed by *reward* and *punishment sensitivity*. Therefore, the effect of decreased dopamine function on the strength of the Pavlovian system may be shadowed by the effects of an age-related decrease in executive functions or instrumental abilities related to structural decline.

The significant age-related increase in the *Pavlovian bias* observed in this study is at odds

with previous aging studies using the same task that either reported no age-related differences in the *Pavlovian bias* (Perosa *et al.*, 2020) or that age-related differences in choice behavior were better accounted for by an instrumental bias instead of a *Pavlovian bias* (de Boer *et al.*, 2019). Taken together, these findings would suggest that there is substantial heterogeneity in the strength of the *Pavlovian bias* that cannot be entirely accounted for by age. Thus, it will be interesting in future work to determine how age-related differences in instrumental biases also influence action-valence learning across the human lifespan.

Finally, some important limitations should be considered on interpretation of these findings. In the present study we can only infer developmental and age-related differences from cross-sectional data. These effects should not be assumed to represent underlying causal relationships, nor can we comment on lifespan trajectories. Future longitudinal studies will be needed to address these questions. With regard to modeling age-related differences in choice behavior, the irreducible noise parameter was unexpectedly markedly reduced in midlife compared to younger and older adults indicating that performance accuracy during this age range was less tightly captured by the winning Pavlovian model. Hence we cannot rule out that the effects in midlife/older adults may in part be related to a selection bias.

Conclusions

Our results demonstrate a dichotomy between prepotent biases that influence learning at either end of the lifespan with a predominant preference for action responses in children/adolescents compared to reduced instrumental learning from both reward and punishment and increased reliance on Pavlovian heuristics in older age. Collectively, our results emphasize the importance of orthogonally manipulating action requirements and outcome valence to further understand instrumental learning capabilities across different stages of the human lifespan. Such characteristics may underline important evolutionary conserved mechanisms i.e. heightened action learning in adolescents necessary to facilitate active exploration and independence into adulthood or alternatively adaptation to maintain decision-making abilities despite declining learning ability in old age.

Acknowledgements

The study was supported by the German Research Foundation SFB 779 (TP A03, TP A07, TP A08) and the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 720270 (HBP SGA1). Research in the authors' labs was also supported by the EU/EFRE-funded "Autonomy In Old Age" Initiative of the State of Saxony-Anhalt. AR was supported by a research grant from the German Research Council (DFG RI 2964-1). MG-M and LdB were supported by a research grant from the Swedish Research Council (VR521-2013-2589) awarded to MG-M. The authors thank Lena Pietro, Carolin Breitling, Anne Hochkeppler, Iris Mann, Catherine Liebeau, Timo Lemme and Maria Watermann for their help during data acquisition. Moreover, we thank Fabian Senner and Marius Keute for their help during preparation of the manuscript.

Disclosure

The authors have no financial or non-financial competing interests to declare.

References

Acosta-Cabronero J, Betts MJ, Cardenas-Blanco A, Yang S, Nestor PJ. In Vivo MRI Mapping of Brain Iron Deposition across the Adult Lifespan. J Neurosci 2016; 36: 364–74.

Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex: one decade on. Trends Cogn Sci 2014; 18: 177–85.

Bäckman L, Nyberg L, Lindenberger U, Li S-C, Farde L. The correlative triad among aging, dopamine, and cognition: Current status and future prospects. Neurosci Biobehav Rev 2006; 30: 791–807.

Blakemore S-J, Robbins TW. Decision-making in the adolescent brain. Nat Neurosci 2012; 15: 1184–91.

de Boer L, Axelsson J, Chowdhury R, Riklund K, Dolan RJ, Nyberg L, et al. Dorsal striatal dopamine D1 receptor availability predicts an instrumental bias in action learning. Proc Natl Acad Sci 2019; 116: 261–70.

de Boer L, Axelsson J, Riklund K, Nyberg L, Dayan P, Bäckman L, et al. Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age. eLife 2017; 6: e26424.

Callaghan MF, Freund P, Draganski B, Anderson E, Cappelletti M, Chowdhury R, et al. Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging. Neurobiol Aging 2014; 35:1862–1872

Casey BJ. Beyond Simple Models of Self-Control to Circuit-Based Accounts of Adolescent Behavior. Annu Rev Psychol 2015; 66: 295–319.

Casey BJ, Getz S, Galvan A. The adolescent brain. Dev Rev 2008; 28: 62–77.

Cavanagh JF, Eisenberg I, Guitart-Masip M, Huys Q, Frank MJ. Frontal Theta Overrides Pavlovian Learning Biases. J Neurosci 2013; 33: 8541–8.

Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, et al. Dopamine restores reward prediction errors in old age. Nat Neurosci 2013; 16: 648–53.

Chowdhury R, Guitart-Masip M, Lambert C, Dolan RJ, Düzel E. Structural integrity of the substantia nigra and subthalamic nucleus predicts flexibility of instrumental learning in older-age individuals. Neurobiol Aging 2013; 34: 2261–70.

Davidow JY, Foerde K, Galván A, Shohamy D. An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. Neuron 2016; 92: 93–9.

Daw, N., Trial-by-trial data analysis using computational models. In *Decision Making, Affect, and Learning*, Delgado, M. R.; Phelps, E. A.; Robbins, T. W., Eds. Oxford University Press: 2011.

Draganski B, Ashburner J, Hutton C, Kherif F, Frackowiak RSJ, Helms G, et al. Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ). NeuroImage 2011; 55: 1423–34.

van Duijvenvoorde ACK, Zanolie K, Rombouts SARB, Raijmakers MEJ, Crone EA. Evaluating the Negative or Valuing the Positive? Neural Mechanisms Supporting Feedback-Based Learning across Development. J Neurosci 2008; 28: 9495–503.

Eppinger B, Kray J, Mock B, Mecklinger A. Better or worse than expected? Aging, learning, and the ERN. Neuropsychologia 2008; 46: 521–39.

Galvan. Adolescent development of the reward system. Front Hum Neurosci 2010; 4:1-9

Galvan A, Hare TA, Parra CE, Penn J, Voss H, Glover G, et al. Earlier Development of the Accumbens Relative to Orbitofrontal Cortex Might Underlie Risk-Taking Behavior in Adolescents. J Neurosci 2006; 26: 6885–92.

Geurts DEM, Huys QJM, den Ouden HEM, Cools R. Aversive Pavlovian Control of Instrumental Behavior in Humans. J Cogn Neurosci 2013; 25: 1428–41.

Guitart-Masip M, Barnes GR, Horner A, Bauer M, Dolan RJ, Duzel E. Synchronization of Medial Temporal Lobe and Prefrontal Rhythms in Human Decision Making. J Neurosci 2013; 33: 442–51.

Guitart-Masip M, Chowdhury R, Sharot T, Dayan P, Duzel E, Dolan RJ. Action controls dopaminergic enhancement of reward representations. Proc Natl Acad Sci 2012; 109: 7511–6.

Guitart-Masip M, Economides M, Huys QJM, Frank MJ, Chowdhury R, Duzel E, et al. Differential, but not opponent, effects of l-DOPA and citalopram on action learning with reward and punishment. Psychopharmacology (Berl) 2014; 231: 955–66.

Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ. Go and no-go learning in reward and punishment: Interactions between affect and effect. NeuroImage 2012; 62: 154–66.

Hämmerer D, Li S-C, Müller V, Lindenberger U. Life Span Differences in Electrophysiological Correlates of Monitoring Gains and Losses during Probabilistic Reinforcement Learning. J Cogn Neurosci 2011; 23: 579–92.

Huys QJM, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, et al. Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. PLoS Comput Biol 2011; 7: e1002028.

Karrer TM, Josef AK, Mata R, Morris ED, Samanez-Larkin GR. Reduced dopamine receptors and transporters but not synthesis capacity in normal aging adults: a meta-analysis. Neurobiol Aging 2017; 57: 36–46.

Kass RE, Raftery AE. Bayes Factors. J Am Stat Assoc 1995; 90: 773.

Kray J, Eber J, Lindenberger U. Age differences in executive functioning across the lifespan: The role of verbalization in task preparation. Acta Psychol (Amst) 2004; 115: 143–65.

Kuo H-Y, Liu F-C. Synaptic Wiring of Corticostriatal Circuits in Basal Ganglia: Insights into the Pathogenesis of Neuropsychiatric Disorders. eneuro 2019; 6: ENEURO.0076-19.2019.

Lindenberger U. Human cognitive aging: Corriger la fortune? Science 2014; 346: 572–8.

Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl) 2007; 191: 507–20.

Palminteri S, Khamassi M, Joffily M, Coricelli G. Contextual modulation of value signals in reward and punishment learning. Nat Commun 2015; 6: 8096.

Palmiter RD. *Dopamine Signaling in the Dorsal Striatum Is Essential for Motivated Behaviors*. Ann N Y Acad Sci 2008; 1129: 35–46.

Peeters M, Oldehinkel T, Vollebergh W. Behavioral Control and Reward Sensitivity in Adolescents' Risk Taking Behavior: A Longitudinal TRAILS Study. Front Psychol 2017; 8:231

Perosa V, de Boer L, Ziegler G, Apostolova I, Buchert R, Metzger C, et al. The Role of the Striatum in Learning to Orthogonalize Action and Valence: A Combined PET and 7 T MRI Aging Study. Cereb Cortex 2020; 30: 3340–51.

Richter A, Guitart-Masip M, Barman A, Libeau C, Behnisch G, Czerney S, et al. Valenced action/inhibition learning in humans is modulated by a genetic variant linked to dopamine D2 receptor expression. Front Syst Neurosci 2014; 8: 140

Rutledge RB, Skandali N, Dayan P, Dolan RJ. Dopaminergic Modulation of Decision Making and Subjective Well-Being. J Neurosci 2015; 35: 9811–22.

Salamone JD, Correa M. The Mysterious Motivational Functions of Mesolimbic Dopamine. Neuron 2012; 76: 470–85.

Samanez-Larkin GR, Gibbs SEB, Khanna K, Nielsen L, Carstensen LL, Knutson B. Anticipation of monetary gain but not loss in healthy older adults. Nat Neurosci 2007; 10: 787–91.

Samanez-Larkin GR, Levens SM, Perry LM, Dougherty RF, Knutson B. Frontostriatal White Matter Integrity Mediates Adult Age Differences in Probabilistic Reward Learning. J Neurosci 2012; 32: 5333–7.

Schott BH, Niehaus L, Wittmann BC, Schutze H, Seidenbecher CI, Heinze H-J, et al. Ageing and early-stage Parkinson's disease affect separable neural mechanisms of mesolimbic reward processing. Brain 2007; 130: 2412–24.

Schultz W. Dopamine signals for reward value and risk: basic and recent data. Behav Brain Funct 2010; 6: 24.

Shulman EP, Smith AR, Silva K, Icenogle G, Duell N, Chein J, et al. The dual systems model: Review, reappraisal, and reaffirmation. Dev Cogn Neurosci 2016; 17: 103–17.

Somerville LH, Hare T, Casey BJ. Frontostriatal Maturation Predicts Cognitive Control Failure to Appetitive Cues in Adolescents. J Cogn Neurosci 2011; 23: 2123–34.

Spear LP. The adolescent brain and age-related behavioral manifestations. Neurosci Biobehav Rev 2000; 24: 417–463.

Steiger TK, Weiskopf N, Bunzeck N. Iron Level and Myelin Content in the Ventral Striatum Predict Memory Performance in the Aging Brain. J Neurosci 2016; 36: 3552–8.

Sutton RS, Barto AG. Reinforcement learning: an introduction. Cambridge, Mass: MIT Press; 1998

Swart JC, Froböse MI, Cook JL, Geurts DE, Frank MJ, Cools R, et al. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in) action. eLife 2017; 6: e22169.

Talmi D, Seymour B, Dayan P, Dolan RJ. Human Pavlovian Instrumental Transfer. J Neurosci 2008; 28: 360–8.

Van Leijenhorst L, Moor BG, Op de Macks ZA, Rombouts SARB, Westenberg PM, Crone EA. Adolescent risky decision-making: Neurocognitive development of reward and control regions. NeuroImage 2010; 51: 345–55.

van de Vijver I, Ridderinkhof KR, Harsay H, Reneman L, Cavanagh JF, Buitenweg JIV, et al. Frontostriatal anatomical connections predict age- and difficulty-related differences in reinforcement learning. Neurobiol Aging 2016; 46: 1–12.

Walker DM, Bell MR, Flores C, Gulley JM, Willing J, Paul MJ. Adolescence and Reward: Making Sense of Neural and Behavioral Changes Amid the Chaos. J Neurosci 2017; 37: 10855–66.

Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York: 2016; ISBN 978-3-319-24277-4
Zelazo PD, Craik, Fergus I. M. FIM, Booth L. Executive function across the life span. Acta Psychol (Amst) 2004; 115: 167–83.

Ziegler G, Ridgway GR, Blakemore S-J, Ashburner J, Penny W. Multivariate dynamical modelling of structural change during development. NeuroImage 2017; 147: 746–62.