Supplemental Material

Summary of Previous Developmental Reversal-Learning Studies

Supplementary tables 1 and 2 provide an overview of the methods (suppl. Table 1) and results (suppl. Table 2) of three previous development reversal-learning studies (Hauser et al., 2015; Javadi et al., 2014; van der Schaaf et al., 2011).

Pubertal Development

Participants aged 8-17 completed the pubertal developmental scale (PDS), a questionnaire that determines pubertal status based on questions about physical development (Petersen et al., 1988). In addition, an hour after the start of the experiment and in-between tasks, participants provided a 1.8 ml saliva sample, which was analyzed for testosterone levels as a marker of pubertal development. The procedure is described in detail in Master et al., 2020. PDS scores and testosterone levels were highly correlated with age for both males and females (suppl. Fig. 1B), making it difficult to assess them separately. We created quantile groups for pubertal measurements similar to age: For PDS scores, we assigned all participants with score 1 to the pre-pubertal group, and divided the remaining participants into tertiles based on score, which we termed "early", "middle", and "late" puberty. Tertiles were defined separately for males and females to assure sex balance within each group (suppl. Fig. 1A, middle row). For testosterone levels, we created quantiles based on testosterone levels, again defining quantiles separately for males and females, using the same method as to create age quartiles.

Similar developmental patterns arose for pubertal development (PDS, testosterone) as seen for age (suppl. Fig. 2, 3, 4). The main difference was at which time peak performance occurred: in the third quantile based on age (13-15 years), but the fourth quantile based on puberty (suppl. Fig. 2). Parameter trajectories also differed slightly: most notably, p and β showed more abrupt changes based on PDS, with steps between mid and late puberty. α_- and p_{reward} showed a drastic step at puberty onset (between "pre" and "early"; suppl. Fig. 3B). In terms of testosterone, parameters α_-, p_{reward}, and p_{switch} showed U-shaped functions similar to age, but minima occurred in the fourth rather than the third quantile (suppl. Fig. 3C). In terms of parameter PCs as well, trajectories were largely similar between pubertal measures and age. Slight differences included a more unique role of pre-pubertal participants, especially for PC2 in terms of PDS and PC3 for testosterone (suppl. Fig. 4).

Table 1. Overview of studies that have used reversal tasks in developing human populations. This table only includes studies which conducted age comparisons (excluding, e.g., Boehme et al., 2017; Dickstein, Finger, Brotman, et al., 2010; Dickstein, Finger, Skup, et al., 2010; Finger et al., 2008; Harms et al., 2018), and which included the adolescent age range (excluding, e.g., Minto de Sousa et al., 2015). The table highlights participant samples, reversal task designs, and RL modeling quality.

Study	Participant age	Task	RL model	RL model quality
Javadi et al., 2014	14-15 (n=260) 20-39 (n=29)	Select one stimulus on each trial Correct: 70% reward, 30% punishment Incorrect: 40% reward, 60% punishment Reversal: 25% per trial after ≥ 4 correct	Adaptive α (3 parameters)	No model comparison No model validation
Hauser et al., 2015	12-16 (n=19) 20-29 (n=17)	Select one stimulus on each trial Correct: 80% reward, 20% punishment Incorrect: 20% reward, 80% punishment Reversal: 6-10 trials after 3 consec. correct	[Positive vs negative], [fact. vs count.-fact.] α	Model comparison (3 models) No model validation
van der Schaaf et al., 2011	10-11 (n=15) 13-14 (n=15) 16-17 (n=15) 20-25 (n=16)	Predict outcome of highlighted stimulus Correct: 100% reward, 0% punishment Incorrect: 0% reward, 100% punishment Reversal: after 4-6 consecutive correct	No computational model	
Ours	8-17 (n=191) 18-26 (n=66) 25-30 (n=55)	Select one stimulus on each trial Correct: 75% reward, 25% punishment Incorrect: 0% reward, 100% punishment Reversal: After 7-15 rewards	Positive, negative, fact., count.-fact.	Extensive model comparison (7 RL & 16 BI models) Extensive model validation

Table 2. Overview of the results of the studies presented in suppl. Table 1. We focus on difference in overall performance between age groups, differences in the number of reversals obtained (another measure of performance because reversals were performance dependent), and difference in RL model parameters. Note that studies were neither comparable in terms of task design (suppl. Table 1), nor in terms of computational models (see original papers for summary), meaning that results are not directly comparable.

Study	Performance	Number of reversals	RL model results
Javadi et al., 2014	No age difference	More in adults	$log(\gamma)$ lower in adolescents Larger RPEs in adolescents after correct responses but negative feedback
Hauser et al., 2015	No age difference	No age difference	$\alpha_{-\ factual}$ higher in adolescents
van der Schaaf et al., 2011	Linear increase non-reversal trials (asymptote in adolescence) Inverse U-shape reversal trials (max in adolescence)	Linear increase with age	No model
Ours	Inverse U-shape asymptotic trials (max in mid-adolescence) Inverse U-shape reversal trials (max in adolescence)	NA	p increases with age, asymptotes in late adolescence β increases with age, asymptotes in late adolescence α_- U-shape, lowest in mid-adolescence α_+ step function, larger in adults

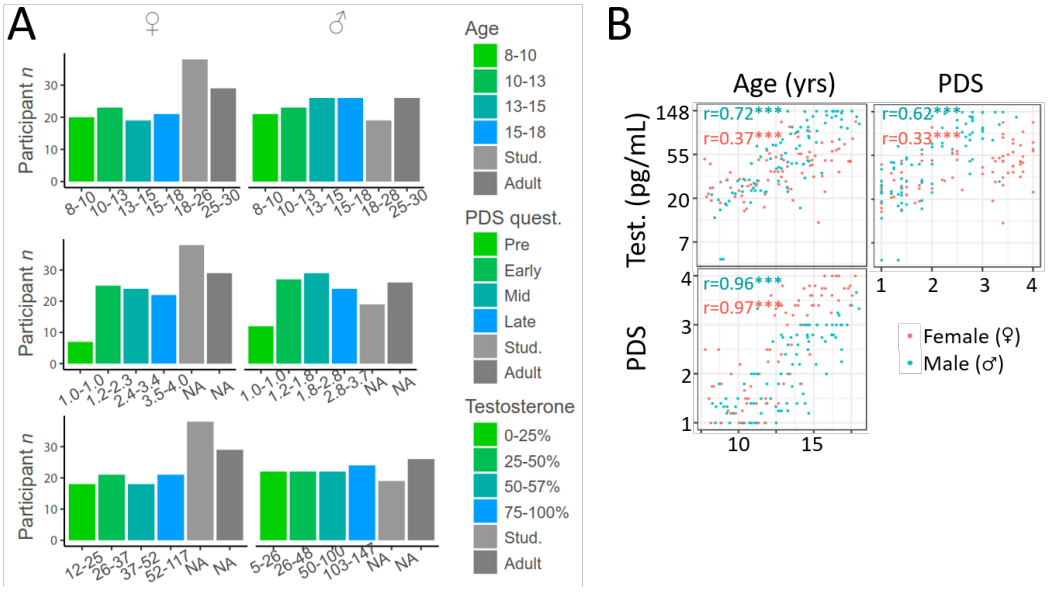

Figure 1. Participant sample and pubertal development. A) Number of participants in each bin, separately for each sex. Top: Age quantiles, which are the basis of Figures 2, 3, 4, and 5, and suppl. Figures 8, 10, 11, and 12. Numbers on the x-axis indicate the age ranges that went into each quantile bin, which differed slightly between males and females. The legend shows the names of the bins used throughout the paper. Middle: Bins based on the pubertal development questionnaire (PDS), which was available only for participants aged 8-17. The numbers on the x-axis show the ranges of each bins, which differed substantially between sexes. The legend shows the bin names after combining males and females. Bottom: Bins based on salivary testosterone levels, using the same conventions as above. B) Correlations between age, testosterone levels (Test.), and PDS questionnaire, for male and female participants aged 8-17. Stars refer to p-values, using the same convention as in main text figures.

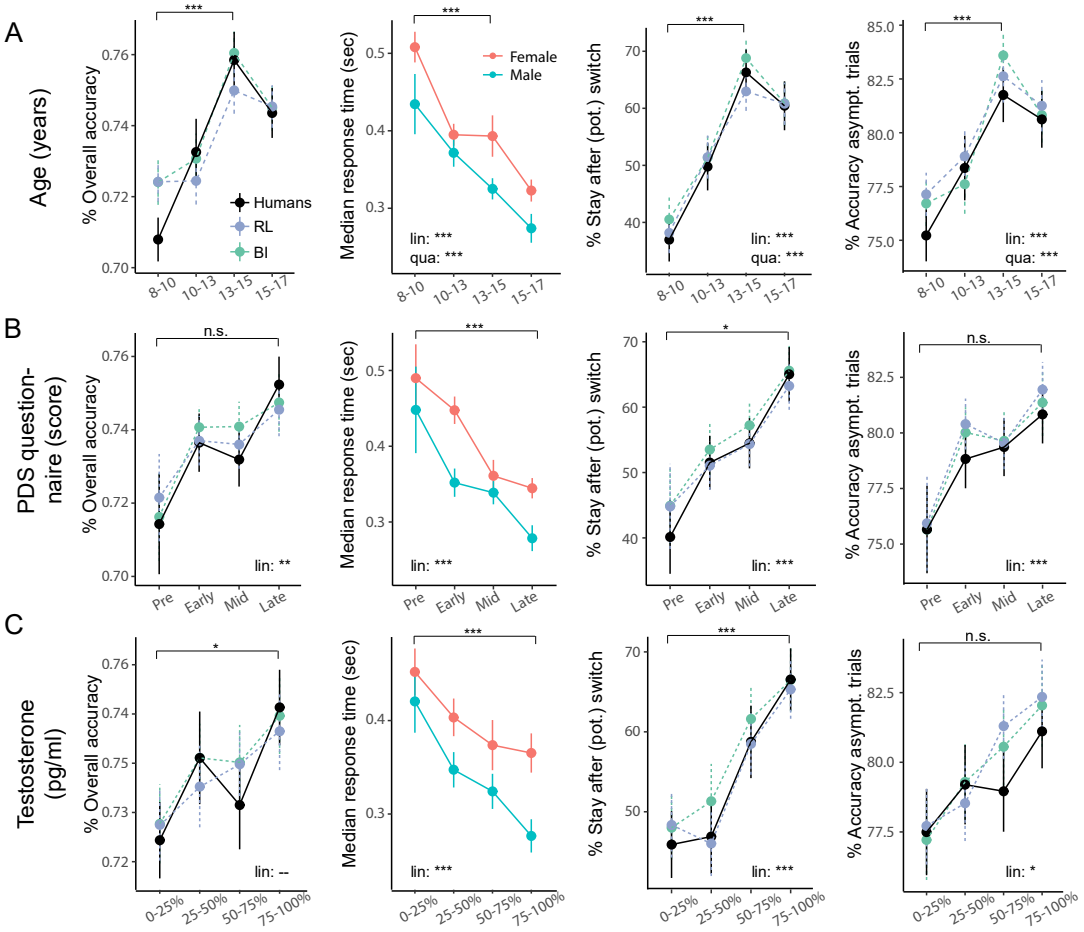

Figure 2. Behavior broken up by age / PDS / Testosterone bins. Significance bars and stars show the results of planned t-tests. A) Same data as in Fig. 3. Planned t-tests compared 8-10 year olds to 13-to-15-year-olds. B) Same measures, but broken up by PDS bins. T-tests compared pre-pubertal to late-pubertal participants. C) Same measures, broken up by testosterone bins. T-tests compared participants in the first quantile in terms of testosterone levels to participants in the fourth quantile.

Table 3. Statistics of mixed-effects regression models predicting performance measures from sex (male, female) and puberty measures (PDS questionnaire / salivary testosterone). Only participants who had these measures were included in the model, restricting it to participants under the age of 18. Overall accuracy, stay after potential (pot.) switch, and asymptotic performance were modeled using logistic regression, and z-scores are reported. Log-transformed response times on correct trials were modeled using linear regression, and t-values are reported. * $p < .05$; ** $p < .01$, *** $p < .001$.

Performance measure (Figure)	Predictor	β	z / t	p	sig.
Overall accuracy (2B, left)	PDS	0.069	2.9	0.0038	**
	Sex	0.017	0.37	0.71	
Response times (2B, 2^{nd}-to-left)	PDS	-0.13	-4.9	< 0.001	***
	Sex	0.25	4.8	< 0.001	***
Stay after (pot.) switch (2B, 2^{nd}-to-right)	PDS	0.48	3.5	< 0.001	***
	Sex	0.76	2.9	0.0036	**
Asymptotic performance (2B, right)	PDS	0.25	4.2	< 0.001	***
	Sex	0.098	0.9	0.39	
Overall accuracy (2C, left)	Test.	< 0.0001	1.2	0.24	
	Sex	0.032	0.69	0.49	
Response times (2C, 2^{nd}-to-left)	Test.	-0.0034	-5.1	< 0.001	***
	Sex	0.010	1.9	0.049	*
Stay after (pot.) switch (2C, 2^{nd}-to-right)	Test.	0.012	3.5	< 0.001	***
	Sex	0.27	1.0	0.29	
Asymptotic performance (2C, right)	Test.	0.0034	2.2	0.029	*
	Sex	0.12	1.0	0.34	

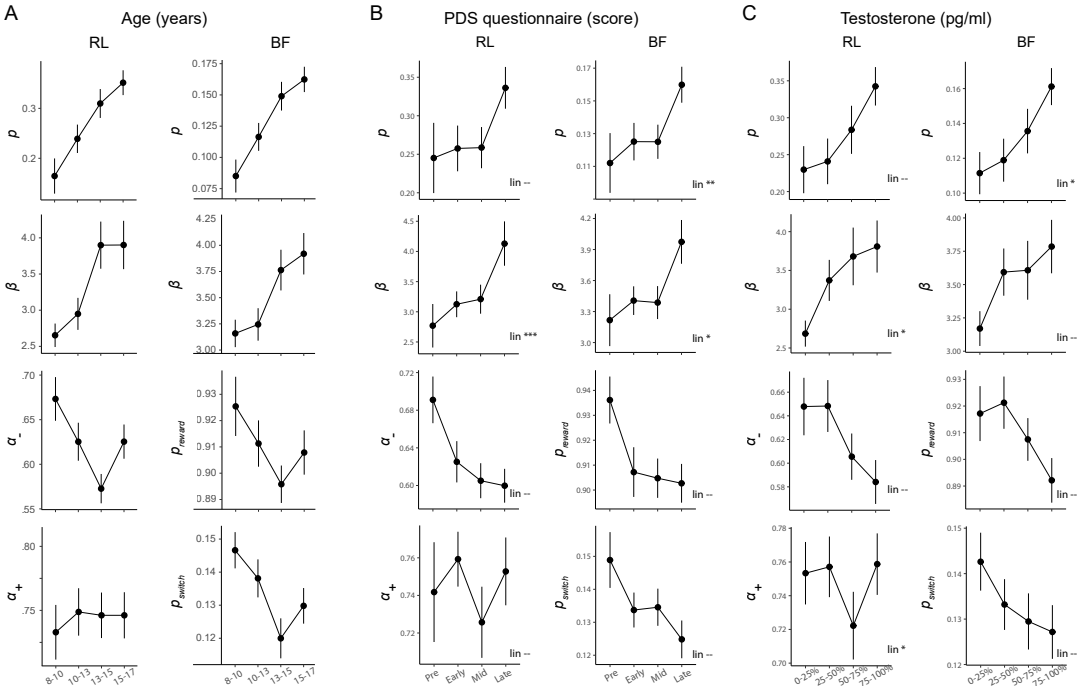

Figure 3. Model parameters broken up by age / PDS / Testosterone bins. A) Participants younger than 18 years of age, reproduced from Fig. 4. B)-C) Same data, broken up by PDS / testosterone bins. "lin." indicates whether a linear effect of the measure of interest (PDS / testosterone) reached significance in a linear regression model.

Table 4. Parameter estimates and statistics from hierarchical model fitting, for pubertal predictors (PDS questionnaire, salivary testosterone), for participants under the age of 18. Significance tests against 0 for parameters whose range includes 0, NA otherwise.

Model	Parameter	$\mu + -sd$	95% CI	p-value	sig.
PDS					
4-param. BI	p_{int}	$0.11 + -0.013$	$[0.082, 0.13]$	< 0.001	***
	p_{sd}	$0.089 + -0.0085$	$[0.073, 0.11]$	0	NA
	p_{lin}	$0.022 + -0.0096$	$[0.0039, 0.041]$	0.0086	**
	β_{int}	$3.81 + -0.26$	$[3.31, 4.34]$	0	NA
	β_{sd}	$1.25 + -0.14$	$[0.98, 1.53]$	0	NA
	β_{lin}	$0.31 + -0.16$	$[-0.018, 0.62]$	0.028	*
	$p_{reward\ int}$	$0.88 + -0.019$	$[0.84, 0.92]$	0	NA
	$p_{reward\ sd}$	$0.060 + -0.011$	$[0.038, 0.082]$	0	NA
	$p_{reward\ lin}$	$< 0.001 + -0.010$	$[-0.019, 0.020]$	0.48	–
	$p_{switch\ int}$	$0.16 + -0.016$	$[0.13, 0.20]$	0	NA
	$p_{switch\ sd}$	$0.067 + -0.0070$	$[0.053, 0.080]$	0	NA
	$p_{switch\ lin}$	$-0.0098 + -0.0099$	$[-0.029, 0.0099]$	0.16	–
4-param. RL	p_{int}	$0.25 + -0.026$	$[0.20, 0.30]$	< 0.001	***
	p_{sd}	$0.24 + -0.019$	$[0.20, 0.28]$	0	NA
	p_{lin}	$0.039 + -0.024$	$[-0.0093, 0.087]$	0.054	–
	β_{int}	$3.15 + -0.13$	$[2.90, 3.41]$	0	NA
	β_{sd}	$1.37 + -0.13$	$[1.12, 1.62]$	0	NA
	β_{lin}	$0.41 + -0.13$	$[0.17, 0.66]$	< 0.001	***
	$\alpha_{-\ int}$	$0.60 + -0.016$	$[0.56, 0.62]$	0	NA
	$\alpha_{-\ sd}$	$0.16 + -0.013$	$[0.14, 0.18]$	0	NA
	$\alpha_{-\ lin}$	$-0.0155 + -0.017$	$[-0.048, 0.019]$	0.18	–
	$\alpha_{+\ int}$	$0.66 + -0.028$	$[0.61, 0.72]$	0	NA
	$\alpha_{+\ sd}$	$0.35 + -0.034$	$[0.023, 0.15]$	0	NA
	$\alpha_{+\ lin}$	$0.0085 + -0.027$	$[-0.048, 0.059]$	0.38	–
Testosterone					
4-param. BI	p_{int}	$0.11 + -0.013$	$[0.081, 0.13]$	< 0.001	***
	p_{sd}	$0.089 + -0.0084$	$[0.073, 0.11]$	0	NA
	p_{lin}	$0.02 + -0.010$	$[0.0023, 0.040]$	0.015	*
	β_{int}	$3.78 + -0.26$	$[3.29, 4.31]$	0	NA
	β_{sd}	$1.28 + -0.14$	$[1.00, 1.55]$	0	NA
	β_{lin}	$0.12 + -0.17$	$[-0.20, 0.45]$	0.22	–
	$p_{reward\ int}$	$0.88 + -0.019$	$[0.85, 0.92]$	0	NA
	$p_{reward\ sd}$	$0.056 + -0.011$	$[0.035, 0.077]$	0	NA
	$p_{reward\ lin}$	$-0.0135 + -0.010$	$[-0.033, 0.0081]$	0.90	–
	$p_{switch\ int}$	$0.16 + -0.016$	$[0.13, 0.19]$	0	NA
	$p_{switch\ sd}$	$0.067 + -0.0069$	$[0.054, 0.081]$	0	NA
	$p_{switch\ lin}$	$-0.0082 + -0.010$	$[-0.029, 0.012]$	0.22	–
4-param. RL	p_{int}	$0.24 + -0.025$	$[0.20, 0.29]$	< 0.001	***
	p_{sd}	$0.24 + -0.0195$	$[0.20, 0.28]$	0	NA
	p_{lin}	$0.038 + -0.025$	$[-0.0091, 0.190]$	0.066	–
	β_{int}	$3.16 + -0.14$	$[2.89, 3.43]$	0	NA
	β_{sd}	$1.42 + -0.13$	$[1.17, 1.69]$	0	NA
	β_{lin}	$0.28 + -0.13$	$[0.037, 0.54]$	0.013	*
	$\alpha_{-\ int}$	$0.60 + -0.017$	$[0.55, 0.62]$	0	NA
	$\alpha_{-\ sd}$	$0.16 + -0.013$	$[0.13, 0.18]$	0	NA
	$\alpha_{-\ lin}$	$-0.035 + -0.018$	$[-0.070, -0.0016]$	0.24	–
	$\alpha_{+\ int}$	$0.66 + -0.028$	$[0.61, 0.72]$	0	NA
	$\alpha_{+\ sd}$	$0.10 + -0.030$	$[0.045, 0.16]$	0	NA
	$\alpha_{+\ lin}$	$-0.017 + -0.026$	$[-0.066, 0.036]$	0.015	*

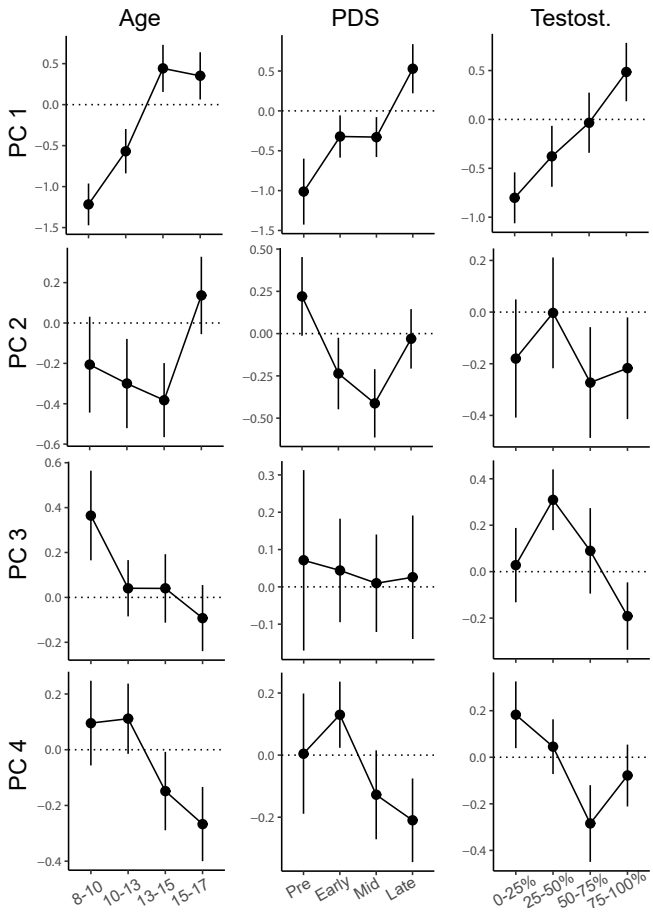

Figure 4. Model parameter PCs broken up by age / PDS / Testosterone bins. Left row: Participants younger than 18 years of age, reproduced from Fig. 5. Middle (right) row: same data, but broken up by PDS (testosterone) bins.

Assessing the Effects of Puberty After Controlling for Age

We next sought to control for age and examine the effect of puberty alone. To this end, we investigated the continuous effects of puberty within each age bin, to eliminate confounds with age (Master et al., 2020). In concordance with the finding that behavior peaked in the third age bin (13-15 years), but in the fourth PDS bin (75-100th percentile), all measures of behavior increased qualitatively with respect to PDS in the third and fourth age bins (suppl. Fig. 5A, right-most column). Nevertheless, this pattern is difficult to interpret because pubertal status was heavily confounded with sex in the fourth age bin, such that girls scored higher on the PDS questionnaire than boys of the same age, in concordance with typical age differences in pubertal maturation. Within the age bins that contained participants across the entire range of pubertal status (10-13, 13-15, and 15-17 years), few significant effects of PDS (suppl. Fig. 5A) or salivary testosterone levels (suppl. Fig. 5B) were observed, possibly including some that occurred by chance. In our data, stay after (pot.) switch trials showed a qualitative decrease with PDS score in 10-13 year olds, was constant in 13-to-15-year-olds, and showed a qualitative increase in 15-17 year olds. This could indicate a weak U-shaped effect or simply experimental noise.

In the case of fit model parameters, pubertal development did not show significant positive relationships with choice parameters p and β, which we might predict if pubertal development was a driving mechanism in growth for these parameters between ages 8-18 (suppl. Table 5; suppl. Fig. 6, 7). In terms of learning parameters, pubertal development also did not show significant negative relationships with α_- and α_+ (RL), or p_{reward} and p_{switch} (BI), which we might predict if pubertal

onset was driving the decrease of these parameters between ages 8-15. If anything, we saw the opposite pattern in males: α_-, p_{reward}, and p_{switch} showed a qualitatively positive relationship with PDS scores (suppl. Fig. 6) and testosterone (suppl. Fig. 7) in the 10-13 year old age group, and a qualitatively negative relationship with PDS in the 13-to-15-year-olds age group. Overwhelmingly, these relationships were not statistically significant.

Trend relationships found within the 13-15 year-old group included a marginal effect of PDS on α_+ (β=0.075, p=0.092), a marginal effect of sex on p_{switch} in the testosterone model (β=0.047, p=0.078), and a significant interaction between sex and testosterone on p_{switch} (β=0.00097, p=0.015; suppl. Table 5). Note that these statistical tests were not corrected for multiple comparisons, making it possible that these results were observed by chance, and should thus be interpreted carefully. The cross-sectional design of our experiment may limit our ability to detect pubertal effects (Kraemer et al., 2000). It is possible that experiments with greater power, longitudinal studies, and studies of hormone manipulation may further inform these largely negative results.

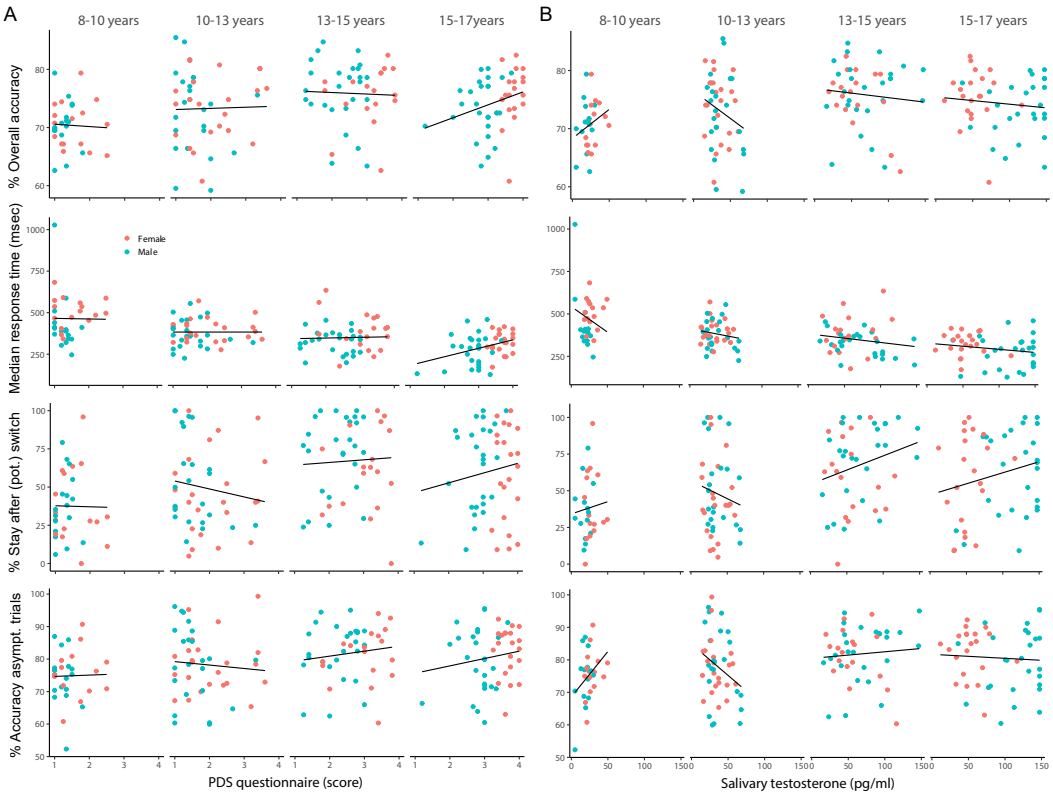

Figure 5. Effects of pubertal status on performance, controlling for age. Each column shows one age group, colors denote sex. Pubertal status was determined by (A) PDS questionnaire, or (B) salivary testosterone.

Table 5. Statistics of regression models testing effects of puberty within the age bin 13-15 years. This bin was chosen because it contained participants across the full range of pubertal development.

Outcome	Predictor	β	p	Sig.
Testosterone				
p (RL)	Test.	-0.00096	0.57	
	Sex	0.062	0.65	
	Interaction	0.0011	0.58	
β (RL)	Test.	-0.022	0.23	
	Sex	1.86	0.22	
	Interaction	0.034	0.13	
α_-	Test.	-0.00033	0.69	
	Sex	0.047	0.48	
	Interaction	0.0014	0.16	
α_+	Test.	-0.00074	0.47	
	Sex	0.0026	0.97	
	Interaction	0.00055	0.65	
p (BF)	Test.	-0.00052	0.43	
	Sex	0.045	0.40	
	Interaction	0.00083	0.30	
β (BF)	Test.	-0.018	0.12	
	Sex	1.12	0.21	
	Interaction	0.021	0.12	
p_{reward}	Test.	-0.00038	0.31	
	Sex	0.0012	0.97	
	Interaction	0.00027	0.54	
p_{switch}	Test.	0.00053	0.10	
	Sex	0.047	0.078	′
	Interaction	0.00097	0.015	*
PDS				
p (RL)	PDS	0.0044	0.95	
	Sex	0.18	0.52	
	Interaction	0.079	0.43	
β (RL)	PDS	0.87	0.30	
	Sex	2.37	0.45	
	Interaction	0.67	0.55	
α_-	PDS	-0.024	0.52	
	Sex	0.071	0.61	
	Interaction	0.063	0.21	
α_+	PDS	0.075	0.092	′
	Sex	0.21	0.21	
	Interaction	0.051	0.39	
p (BF)	PDS	0.011	0.69	
	Sex	0.084	0.45	
	Interaction	0.032	0.43	
β (BF)	PDS	0.62	0.21	
	Sex	1.96	0.30	
	Interaction	0.64	0.34	
p_{reward}	PDS	-0.0080	0.63	
	Sex	0.023	0.72	
	Interaction	0.022	0.33	
p_{switch}	PDS	-0.010	0.51	
	Sex	0.010	0.86	
	Interaction	0.0057	0.82	

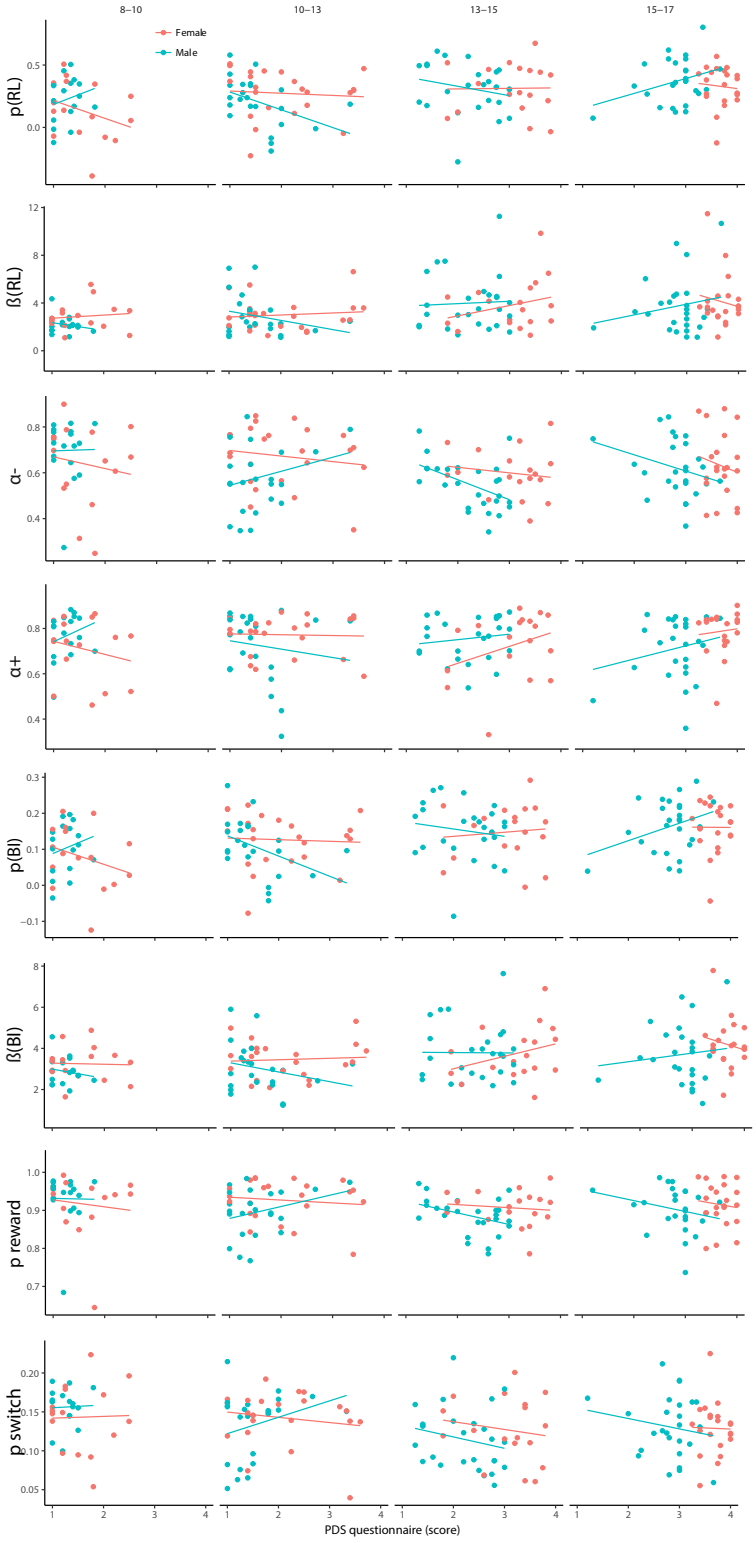

Figure 6. Effects of pubertal status (PDS questionnaire) on model parameters, controlling for age. Each column shows one age group, and colors denote sex.

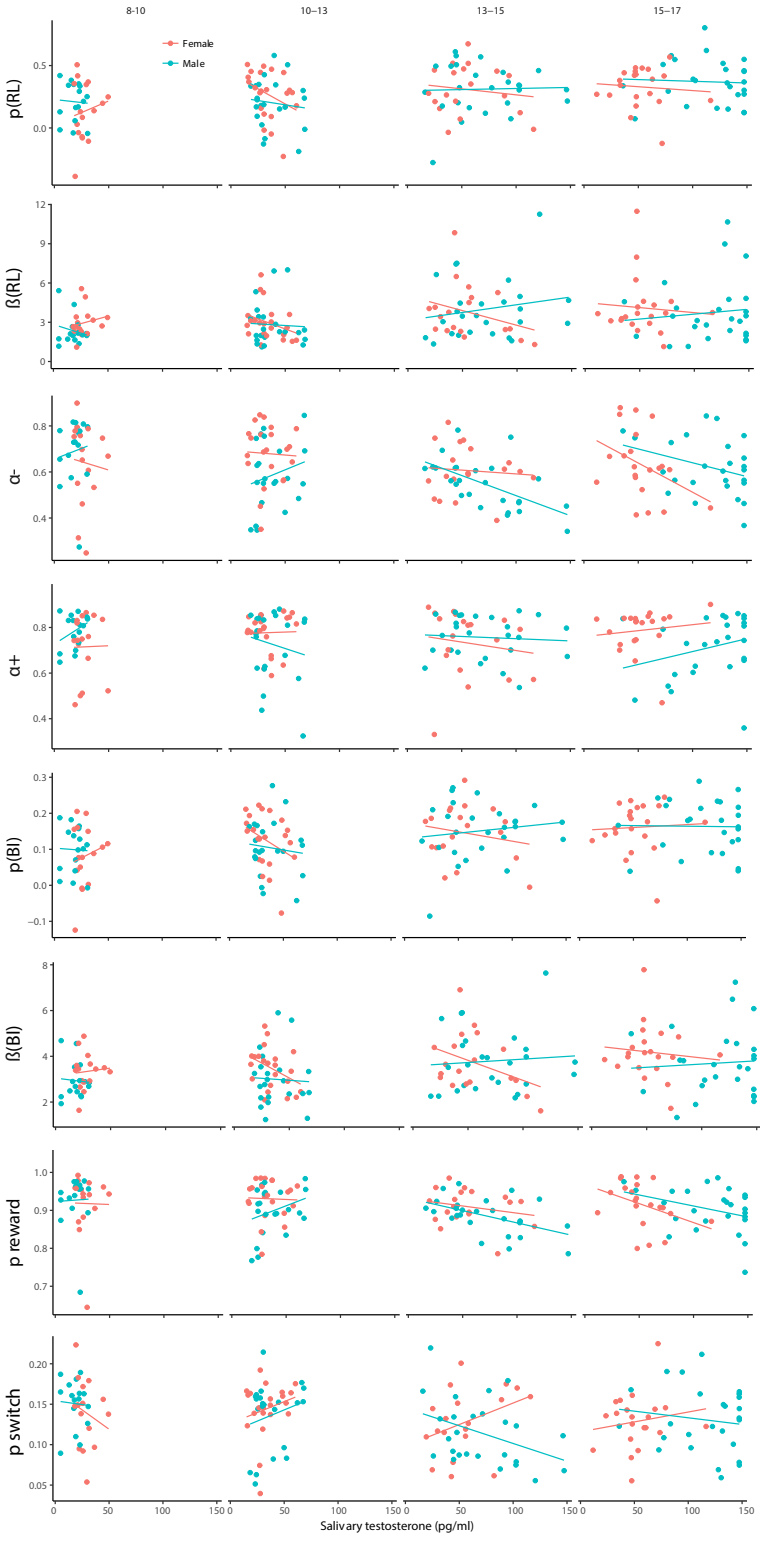

Figure 7. Effects of pubertal status (salivary testosterone levels) on parameters, controlling for age. Each columns shows one parameter. each row one age group, and colors denote sex.

Additional Behavioral Analyses

The youngest children showed the lowest overall and asymptotic accuracy (Fig. 3C, F) and were the most likely to switch after a single negative outcome (Fig. 3E, suppl. Fig. 10B, middle). This explains why they were also fastest at switching (suppl. Fig. 8A, D; suppl. Table 6). Response times were the only performance measure in which 13-to-15-year-olds were outperformed by another age group, university undergraduates (age 18-28; Fig. 2B, 3D). Potential reasons for undergraduates' faster responses include greater familiarity with lab-based psychological experiments, more experience with computers, and increased motivation to finish the task quickly.

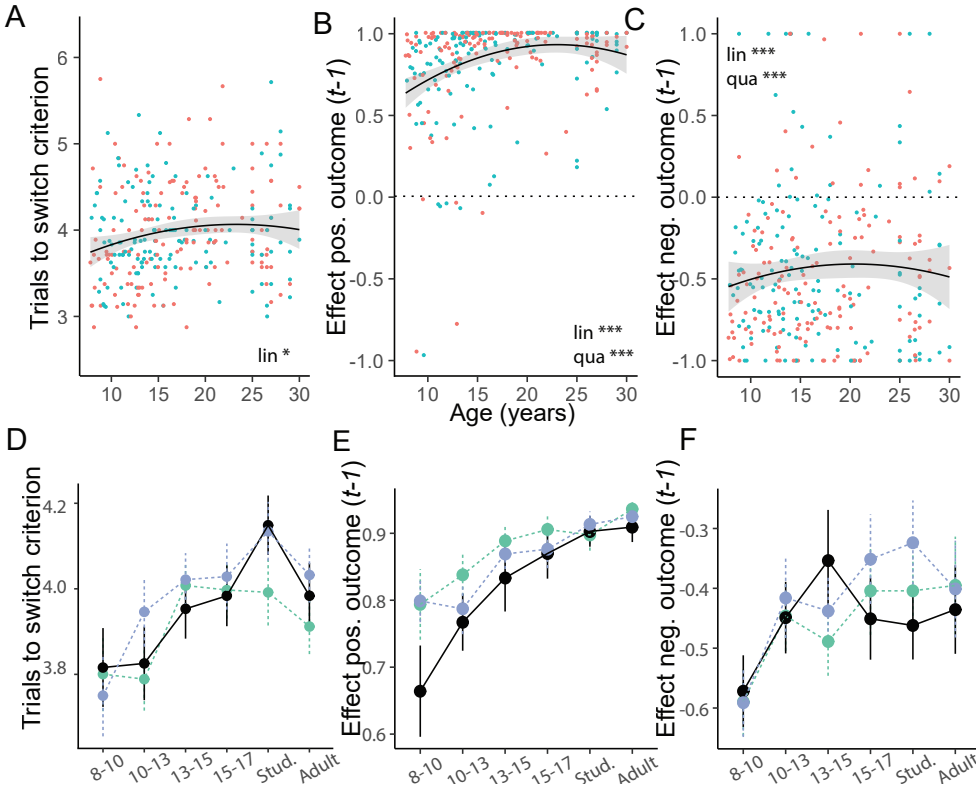

Figure 8. Human behavior (A-C) and model validation (D-F) for additional behavioral measures. (A, D): Number of trials after task switch until participants reached performance criterion (2 correct responses). (B-F): Effect of previous negative (B, E) and positive (C, F) outcomes on choices. "$t-1$": Outcome occurred 1 trial before choice, i.e., delay $i = 1$. Regression weights were tanh transformed for visualization.

Table 6. Statistics of mixed-effects regression model predicting switch criterion from sex (male, female), age (years and months; "lin."), and squared age ("qua."). * $p < .05$; ** $p < .01$, *** $p < .001$.

Behavioral measure (Figure)	Predictor	β	t	p	sig.
Switch criterion (8A)	Age (lin.)	0.067	2.0	0.048	*
	Age (qua.)	-0.0014	-1.6	0.11	
	Trial	0.0059	10.0	< 0.001	***
	Sex	0.0022	0.04	0.97	

Statistics for Regression Models

We conducted regression models predicting future choice from past choice and outcomes. The full statistics of these models are shown in suppl. Table 7.

Table 7. Logistic mixed-effect regression, predicting future actions from past actions and outcomes (methods). The number of predictors ($i \leq 8$) was chosen as to provide the best model fit: $AIC_{i\leq 3}$: 31.046; $AIC_{i\leq 4}$: 31.013; $AIC_{i\leq 5}$: 31.001; $AIC_{i\leq 6}$: 30.981; $AIC_{i\leq 7}$: 30.963; $AIC_{i\leq 8}$**: 30.962**; $AIC_{i\leq 9}$: 30.966; $AIC_{i\leq 10}$: 30.964.

Predictor	**delay** i	β	z	p	**Sig.**
Intercept		-0.01	-0.74	0.46	
Main effects					
Age (lin.)		-0.13	-1.40	0.16	
Age (qua.)		0.12	1.30	0.19	
Pos. outcome	1	2.19	68.09	< 0.001	***
	2	0.84	27.36	< 0.001	***
	3	0.24	7.87	< 0.001	***
	4	0.13	4.30	< 0.001	***
	5	-0.017	-0.54	0.58725	
	6	-0.017	-0.56	0.57548	
	7	-0.0035	-0.12	0.90613	
	8	-0.077	-2.77	0.0057	**
Neg. outcome	1	-0.73	-37.09	< 0.001	***
	2	-0.24	-10.64	< 0.001	***
	3	0.0055	0.22	0.82278	
	4	0.13	5.39	< 0.001	***
	5	0.12	4.87	< 0.001	***
	6	0.12	4.73	< 0.001	***
	7	0.13	5.32	< 0.001	***
	8	0.016	0.71	0.47857	
Interaction age (lin.)					
Pos. outcome	1	0.90	4.50	< 0.001	***
	2	0.84	4.19	< 0.001	***
	3	0.50	2.52	0.012	*
	4	-0.069	-0.35	0.73	
	5	0.088	0.44	0.66	
	6	-0.38	-1.94	0.052	
	7	-0.18	-0.94	0.35	
	8	-0.27	-1.49	0.14	
Neg. outcome	1	0.67	5.27	< 0.001	***
	2	-0.37	-2.48	0.013	*
	3	0.16	1.03	0.30	
	4	-0.089	-0.55	0.58	
	5	0.012	0.07	0.94	
	6	0.066	0.41	0.68	
	7	0.011	0.07	0.94	
	8	-0.068	-0.47	0.63	
Interaction age (qua.)					
Pos. outcome	1	-0.64	-3.14	0.0017	**
	2	-0.89	-4.41	< 0.001	***
	3	-0.38	-1.90	0.057	
	4	0.0020	0.01	0.99	
	5	-0.066	-0.33	0.74	
	6	0.36	1.80	0.072	
	7	0.15	0.75	0.456	
	8	0.29	1.62	0.11	
Neg. outcome	1	-0.56	-4.34	< 0.001	***
	2	0.30	2.00	0.046	*
	3	-0.16	-0.97	0.33	
	4	0.092	0.57	0.57	
	5	-0.0070	-0.04	0.97	
	6	-0.092	-0.57	0.57	
	7	-0.057	-0.35	0.72	
	8	0.064	0.44	0.66	

Meta-Priors for Hierarchical Bayesian Models

Priors for the hierarchical Bayesian models (Fig. 3B and age-less version) were chosen to be uninformative (see suppl. Table 8).

Statistics for Hierarchical Bayesian Models

We verified convergence of the Hierarchical Bayesian Model (Fig. 3B) using the Markov-Chain error, effective sample size (n), and the R-hat statistic ($\hat{R}$), using the functions provided by the PyMC3 toolbox (suppl. Table 9; Salvatier et al., 2016).

Table 9. Statistics for hierarchical Bayesian models. We report the average and the range (min and max over all model parameters) for the two winning models.

Model		MC error	Effective n	$\hat{R}$
4-param. RL	mean	< 0.001	$2,517$	1.001
	range	$[< 0.001; 0.002]$	$[155; 4,261]$	$[1.000; 1.015]$
4-param. BI	mean	0.002	816	1.001
	range	$[< 0.001; 0.01]$	$[281; 1,576]$	$[1.000; 1.004]$

Assessing Model Identifiability using the Generate-and-Recover Precedure

All model fits are relative. In other words, when model A fits data better than model B, there is no guarantee that model A fits the data "well". Both models could fit the data poorly, with model B being even worse than model A. To ensure that our models fit well, we validated our parameter fitting and model comparison method by first simulating and then recovering parameters from each model (Palminteri et al., 2017; Wilson and Collins, 2019). An identifiable model will recover the simulated parameters well during fitting, whereas an unidentifiable model will not. We also compared the results of maximum likelihood and hierarchical Bayesian model fitting using this procedure. Both BF and RL model parameters were recovered well when using hierarchical Bayesian model fitting (age-free model), but recovery was much worse when using maximum likelihood (suppl. Fig. 9A), a well-known fact (Katahira, 2016). Hierarchical Bayesian model fitting also led to more consistent estimates of parameters β and p between both models (suppl. Fig. 9B), showing that this method was especially suited in our case. These results lend credence to the superior fit that can be achieved using Hierarchical Bayesian methods, and to the precision with which model parameter can be estimated.

Table 8. Hyper-priors and priors used in hierarchical Bayesian model fitting. In the age-based model, individuals' parameters were drawn from a Normal distribution around a parameter-specific, age-specific mean θ_m, with parameter-specific standard deviation θ_{sd} (top row of the table; see Fig. 3B for details). In the age-free model, individuals' parameters were drawn from parameter-specific group-level prior distributions (subsequent rows in the table). The shapes of these distributions were based on allowed parameter ranges (e.g., Gamma distribution for parameters with range [0, ∞], Beta distribution for parameters with range [0, 1]). The same prior distribution was used for all individuals, i.e., no age information was present in the age-free model. The distributions of individuals' parameters were themselves parameterized by prior parameters. In the age-based model, prior parameter θ_{sd} was distributed according to a HalfNormal (Normal, truncated at 0; middle section of the table), and parameterized by hyper-parameter $sd = 10$ to allow for a wide, non-informative shape (bottom section). Group-level prior θ_m was defined as an age-based regression function, parameterized by θ_{int}, θ_{lin}, and θ_{qua} for each parameter θ (middle section). The prior on the intercept θ_{int} of each parameter in the age-based model (middle section) had the same shape as the group-level prior distribution in the age-free model (top section), and was parameterized by the same hyper-priors (bottom section). In the age-less model, prior parameters parameterized the distributions of individual model parameters (middle section).

Level	Parameter	Distribution / Value
Shared hyperpriors		
	a	1
	b	1
	m	0
	sd	10
Age-less model		
Parameter priors		
	a_β, b_β, $a_{\alpha+}$, $b_{\alpha+}$, $a_{\alpha-}$, $b_{\alpha-}$, $a_{p\ reward}$, $b_{p\ reward}$, $a_{p\ switch}$, $b_{p\ switch}$	Gamma($\alpha = a$, $\beta = b$)
	m_p	Normal($\mu = m$, $\sigma = sd$)
	sd_p	HalfNormal($\mu = m$, $= sd$)
Indiv. parameters		
	β	Gamma($\alpha = a_\beta$, $\beta = b_\beta$)
	p	Normal($\mu = m_p$, $\sigma = sd_p$)
	α_+	Beta($\alpha = a_{\alpha+}$, $\beta = b_{\alpha+}$)
	α_-	Beta($\alpha = a_{\alpha-}$, $\beta = b_{\alpha-}$)
	p_{reward}	Beta($\alpha = a_{\ reward}$, $\beta = b_{\ reward}$)
	p_{switch}	Beta($\alpha = a_{\ switch}$, $\beta = b_{\ switch}$)
Age-based model		
Parameter priors		
	θ_{sd}, for any parameter θ	HalfNormal($\mu = m$, $= sd$)
	θ_m, for any parameter θ	$\theta_{int} + \theta_{lin}\ age + \theta_{qua}\ age^2$
	β_{int}	Gamma($\alpha = a$, $\beta = b$)
	p_{int}	Normal($\mu = m$, $\sigma = sd$)
	$\alpha_{+\ int}$, $\alpha_{-\ int}$, $p_{reward\ int}$, $p_{switch\ int}$	Beta($\alpha = a$, $\beta = b$)
	θ_{lin}, θ_{qua}, for any parameter θ	Normal($\mu = m$, $\sigma = sd$)
Indiv. parameters		
	θ	Normal($\mu = \theta_m$, $\sigma = \theta_{sd}$)

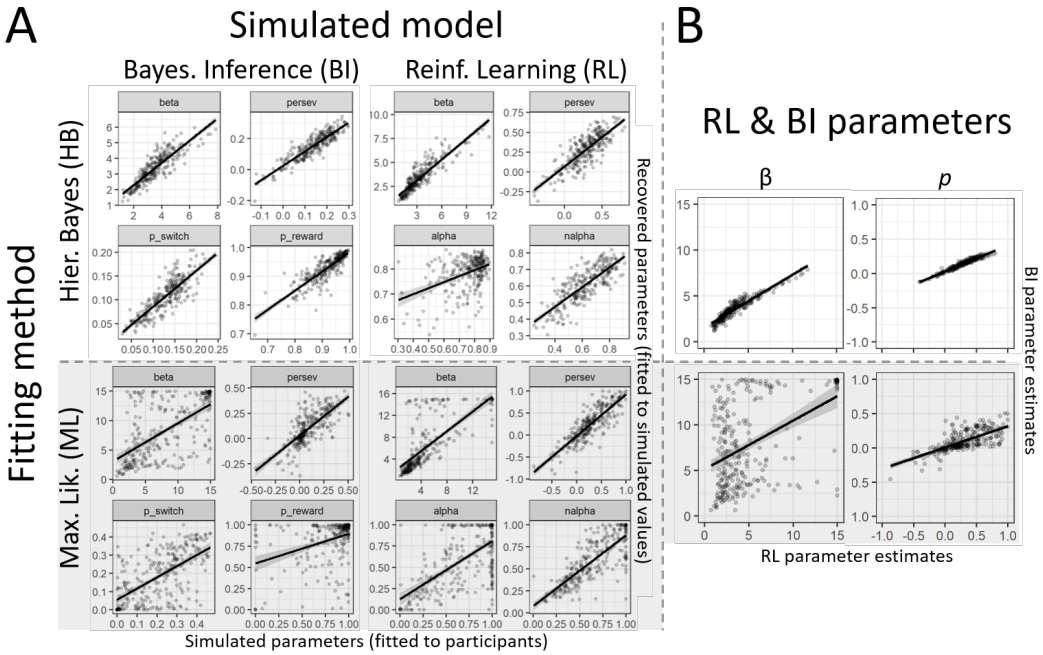

Figure 9. Model validation using hierarchical Bayesian model fitting (top, unshaded), as well as classical Maximum likelihood fitting (bottom, shaded). The results of hierarchical Bayesian fitting are presented in the main text. A) Simulate-and-recover procedure. The x-axes of all graphs show the parameter values of simulated datasets; the y-axes show the recovered parameters obtained by fitting these datasets using the same models. Recovered parameters should be as close to the simulated ones as possible, i.e., lie on the identity line. Black lines and shaded areas indicate best-fit regression lines. The left half presents simulate-and-recover results for the BI model, the right for the RL model. The top half shows the results of hierarchical Bayesian model fitting (our method), the bottom of the standard maximum likelihood method. This figure shows the well-established finding that hierarchical Bayesian model fitting outperformed maximum likelihood. B) Consistency in the estimation of parameters β and p. Human data was fit using RL and BI models to compare the estimates of β (left row) and p (right row) between models. When both (independent) models lead to the same estimates, dots lie on the identity line. This was indeed the case for hierarchical Bayesian fitting (top row), but not for maximum likelihood fitting (bottom row).

Qualitative Fit of RL and BI Models

To test the qualitative fit of our models, we simulated behavior using fitted parameters (from the age-free model; section **??**), and checked whether the simulated behavior was able to reproduce the patterns of interest in the human data (Blohm et al., 2020; Palminteri et al., 2017; Wilson and Collins, 2019). Indeed, both the winning RL and BI models captured human learning curves, as well as sex and age differences, very closely (suppl. Fig. 10). Simpler, non-winning models, on the other hand, failed to capture human characteristics (suppl. Fig. 12, 11).

Raw fitted parameters, obtained from the "age-free" model (Main paper, Methods; suppl. Fig. 12, 11), showed age differences even though age slopes were not part of the fitting model, i.e., individual parameters were not biased by age effects at the group level. To asses effects of age groups, we tested differences in posterior samples of the age-free model. Statistics are shown in suppl. Table 11. To evaluate continuous age effects in a statistically sound way, we used a hierarchical Bayesian model that explicitly modeled age effects (the "age-based" model; Fig. 3B). Significant effects (suppl. Table 10) are shown as lines in suppl. Figures 12 and 11.

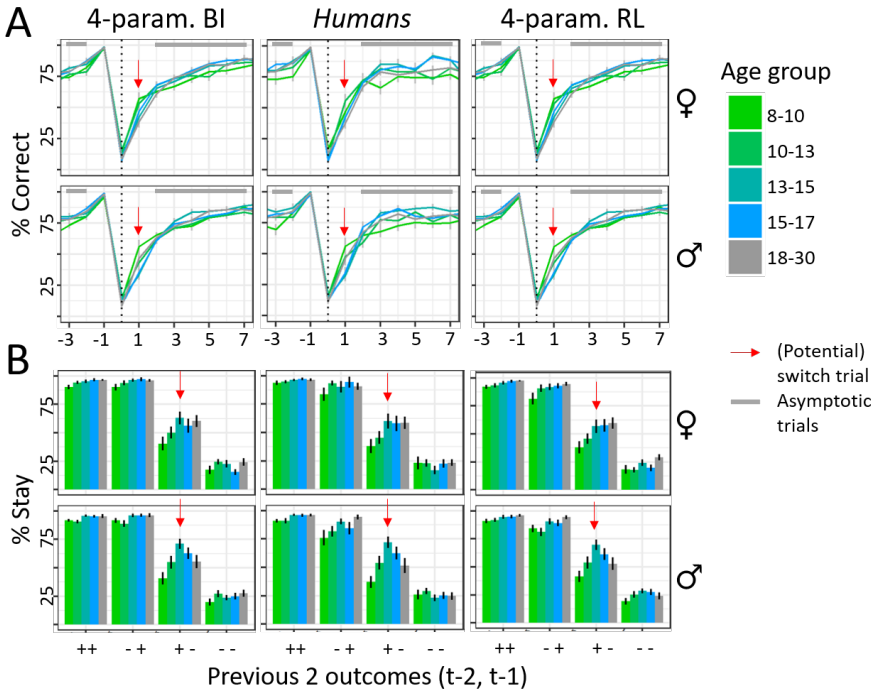

Figure 10. Human and model behavior, showing that models closely reproduced human patterns, A) Behavior in response to switch trials. Colors refer to age groups, red arrows show switch trials, grey bars trials of asymptotic performance. Both models captured quicker switching on switch trials in younger (light green) compared to older participants (blue and grey), and best performance on asymptotic trials in adolescents (green-blue). B) Stay probability in response to outcomes 2 trials back. Both RL and BI replicated human behavior and age differences, including linear increase in staying after positive outcomes ("+ +" and "- +"), and the inverse-U shape on potential switch trials (red arrow; "+ -" condition). Qualitative (non-significant) sex differences were also captured.

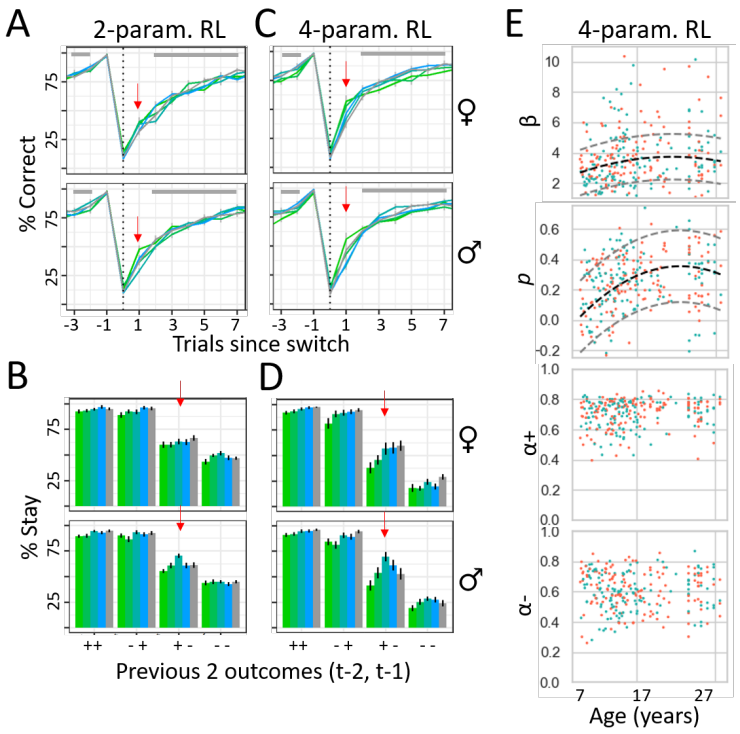

Figure 11. Qualitative fit of different versions of the RL model. Model behavior is shown in the same way as human behavior in suppl. Fig. 10. A-B) Behavior of simulations from the basic, 2-parameter version, with free parameters α and β. Lacking counterfactual updating and the ability to differentiate positive and negative outcomes, the model was unable to capture the shape of human learning curves and age differences. C-D) Behavior of simulations from the winning, 4-parameter version of the RL model, in which free parameters β, p, α_+, and α_- were fitted to participants using hierarchical Bayesian model fitting. To avoid double-dipping into age differences when visualizing the model, we fitted the model *without* access to participants' age (Methods). E) Fitted parameters of each individual, based on the same model. Dashed lines show age differences when significant (Table 9), based on the model with access to participants' age (Fig. 3B). This is the same data as summarized in Fig. 4A-D. Colors denote age groups, red arrow (potential) switch trials, and grey bars asymptotic trials, as in suppl. Fig. 10.

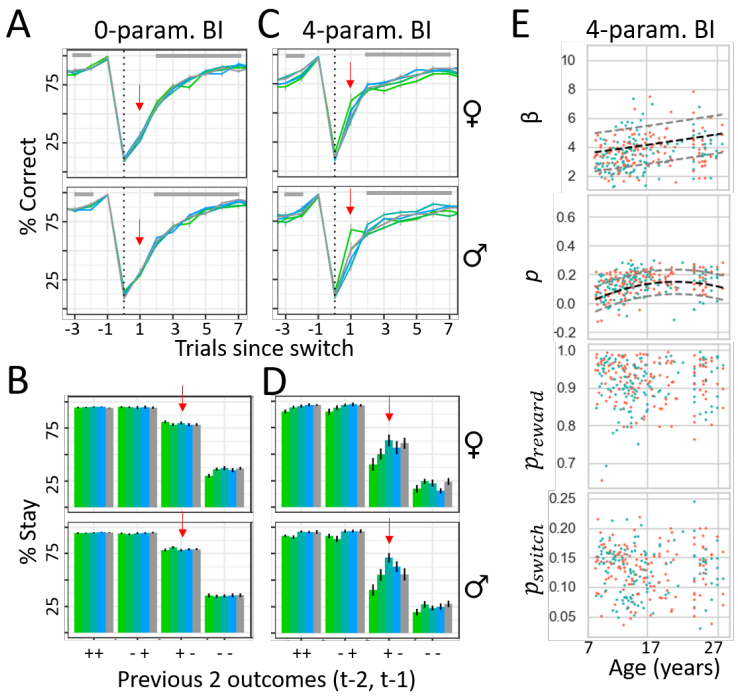

Figure 12. Qualitative fit of different versions of the BI model. Model behavior is shown in the same way as human behavior in suppl. Fig. 10. A-B) Behavior of simulations from the basic, 0-parameter version, in which truthfully $p_{reward} = 0.75$ and $p_{switch} = 0.05$. Lacking free parameters, the model predicted the same behavior for all participants, and was unable to capture age differences. C-D) Behavior of simulations from the winning, 4-parameter version of the BI model, in which free parameters β, p, p_{reward}, and p_{switch} were fitted to participants using hierarchical Bayesian model fitting. To avoid double-dipping into age differences when visualizing the model, we fitted the model *without* access to participants' age (Methods). E) Fitted parameters of each individual, based on the same model. Dashed lines show age differences when significant (suppl. Table 9), based on the model with access to participants' age (Fig. 3B). This is the same data as summarized in Fig. 4E-H.

Table 10. Parameter estimates and statistics from hierarchical model fitting. Significance tests against 0 for parameters whose ranges include 0, NA otherwise.

Model	Parameter	$\mu + -sd$	95% CI	p-value	sig.
4-param. RL	p_{int}	$0.34 + -0.027$	$[0.29, 0.39]$	< 0.001	***
	p_{sd}	$0.24 + -0.015$	$[0.21, 0.26]$	0	NA
	p_{lin}	$0.11 + -0.020$	$[0.075, 0.15]$	< 0.01	**
	p_{qua}	$-0.050 + -0.020$	$[-0.089, -0.012]$	0.0051	**
	β_{int}	$3.48 + -0.15$	$[3.18, 3.79]$	0	NA
	β_{sd}	$1.48 + -0.10$	$[1.29, 1.69]$	0	NA
	β_{lin}	$0.36 + -0.11$	$[0.14, 0.57]$	< 0.001	***
	β_{qua}	$-0.22 + -0.11$	$[-0.42, -0.015]$	0.020	*
	$\alpha_{- int}$	$0.60 + -0.018$	$[0.56, 0.63]$	0	NA
	$\alpha_{- sd}$	$0.16 + -0.0093$	$[0.14, 0.18]$	0	NA
	$\alpha_{- lin}$	$0.011 + -0.015$	$[-0.017, 0.040]$	0.77	
	$\alpha_{- qua}$	$0.013 + -0.014$	$[-0.013, 0.040]$	0.84	
	$\alpha_{+ int}$	$0.73 + -0.034$	$[0.66, 0.79]$	0	NA
	$\alpha_{+ sd}$	$0.081 + -0.021$	$[0.042, 0.12]$	0	NA
	$\alpha_{+ lin}$	$0.055 + -0.024$	$[0.0045, 0.10]$	0.015	*
	$\alpha_{+ qua}$	$-0.015 + -0.021$	$[-0.055, 0.027]$	0.25	
4-param. BI	p_{int}	$0.13 + -0.013$	$[0.11, 0.16]$	< 0.001	***
	p_{sd}	$0.081 + -0.0061$	$[0.069, 0.093]$	0	NA
	p_{lin}	$0.04 + -0.008$	$[0.023, 0.054]$	< 0.001	***
	p_{qua}	$-0.02 + -0.007$	$[-0.038, -0.010]$	< 0.001	***
	β_{int}	$4.27 + -0.27$	$[3.76, 4.83]$	0	NA
	β_{sd}	$1.39 + -0.12$	$[1.16, 1.64]$	0	NA
	β_{lin}	$0.39 + -0.17$	$[0.054, 0.72]$	0.011	*
	β_{qua}	$< 0.001 + -0.16$	$[-0.32, 0.30]$	0.49	
	$p_{reward\ int}$	$0.87 + -0.016$	$[0.84, 0.91]$	0	NA
	$p_{reward\ sd}$	$0.064 + -0.0087$	$[0.046, 0.081]$	0	NA
	$p_{reward\ lin}$	$0.0045 + -0.0096$	$[-0.014, 0.024]$	0.68	
	$p_{reward\ qua}$	$-0.0017 + -0.0085$	$[-0.018, 0.015]$	0.43	
	$p_{switch\ int}$	$0.16 + -0.014$	$[0.14, 0.19]$	0	NA
	$p_{switch\ sd}$	$0.071 + -0.0053$	$[0.062, 0.083]$	0	NA
	$p_{switch\ lin}$	$-0.0066 + -0.0095$	$[-0.025, 0.012]$	0.24	
	$p_{switch\ qua}$	$0.014 + -0.0082$	$[-0.0013, 0.030]$	0.042	*

Table 11. Parameter differences between specific age groups. p-values were obtained by assessing means for each parameter for three age groups (8-10, 13-15, and 18-30) and show in how many MCMC samples the group mean of 8-10 year olds (18-30 year olds) was smaller than the group mean of 13-to-15-year-olds.

Parameter	Compared groups	p-value	sig.
α_-	8-10 vs 13-15	0	***
	13-15 vs Adult	0.0045	**
p_{reward}	8-10 vs 13-15	0.019	*
	13-15 vs Adult	0.078	'
p_{switch}	8-10 vs 13-15	0.023	*
	13-15 vs Adult	0.13	

Using Model Simulations to Elucidate the Role of each PC

We simulated data from our computational models based on the obtained principal components (PCs) in order to visualize the role of each PC. It is common practice to simulate data based on small or large values of a parameter (e.g., smaller or larger decision noise β) to assess the role of this parameter for model behavior (e.g., better or worse performance). We similarly simulated data based on smaller or larger values of each PC to clarify the precise of each PC: We calculated two sets of parameters for each PC, one that represented high levels of this PC ("plus"), and one that represented low values ("minus"). Low levels were determined by subtracting 4 times the inverse-z-scored factor loading of a PC (suppl. Fig. 13, center) from the population mean of each

parameter; low levels were determined by adding it (suppl. Table 12). (For PC2 of the BI model, we added and subtracted 2 times the factor loading instead, to ensure $p_{reward} < 1$.) We then simulated behavior based on the resulting parameters to assess the effect of low versus high values of each PC (suppl. Fig. 13).

Table 12. Parameters used to simulate data to visualize the role of each PC.

	p (RL)	β (RL)	α_-	α_+	p (BI)	β (BI)	p_{reward}	p_{switch}
PC1 (plus)	0.57	6.95	0.45	0.87	0.26	5.67	0.84	0.07
PC1 (minus)	0.04	0.1	0.8	0.65	0.02	1.72	0.98	0.2
PC2 (plus)	0.06	2.65	0.31	0.64	0.1	2.98	0.84	0.12
PC2 (minus)	0.54	4.41	0.94	0.89	0.18	4.41	0.98	0.15
PC3 (plus)	0.76	0.49	0.57	0.74	0.29	1.87	0.85	0.19
PC3 (minus)	-0.16	6.56	0.68	0.78	-0.01	5.52	0.97	0.08
PC4 (plus)	0.15	1.68	0.58	1.19	0.1	3.06	0.88	0.14
PC4 (minus)	0.45	5.38	0.67	0.33	0.18	4.33	0.94	0.13
Parameter mean	0.3	3.53	0.62	0.76	0.14	3.69	0.91	0.13

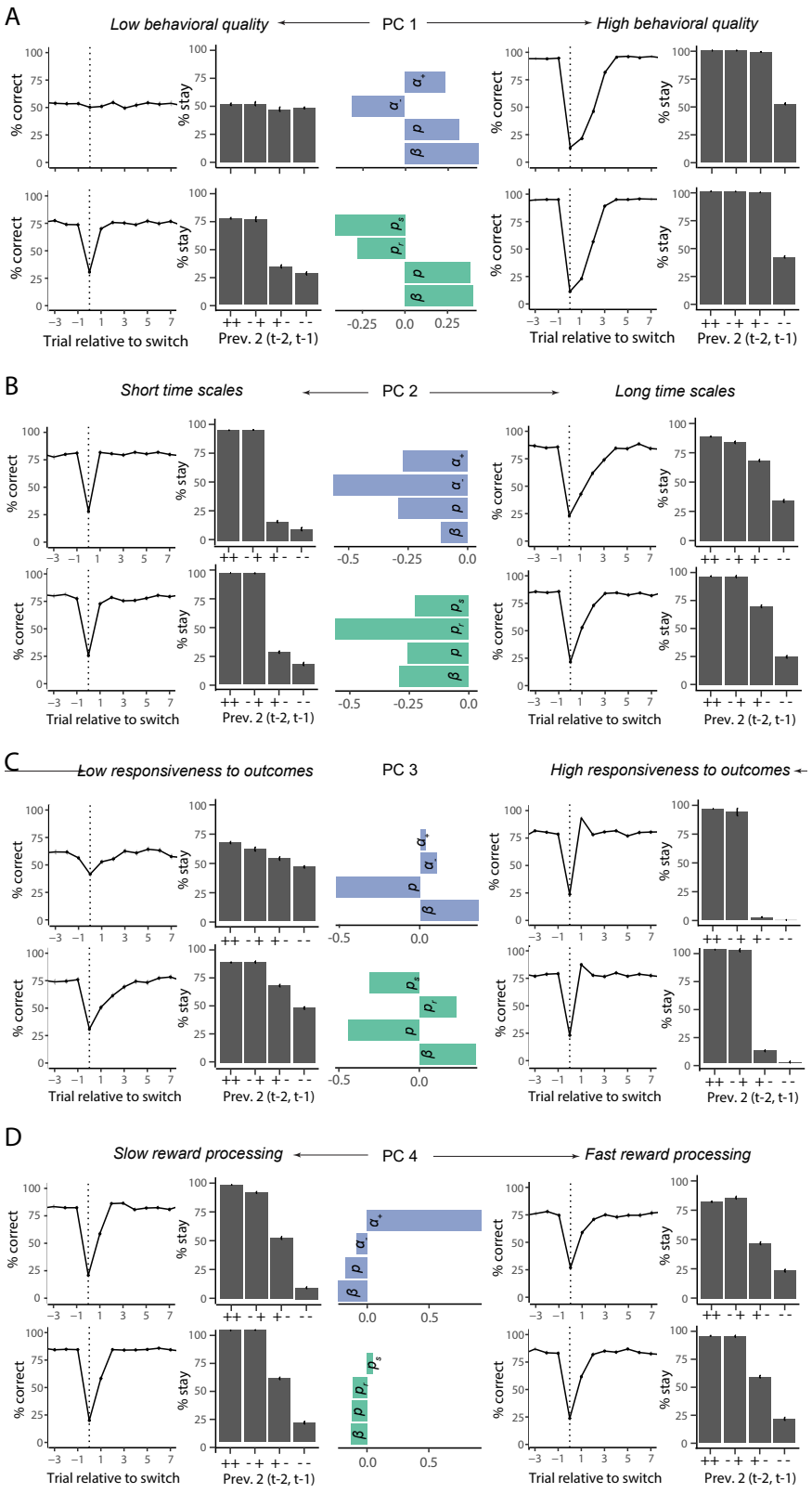

Figure 13. Role of each PC, assessed by simulating behavior at the extremes of each PC. A-D) PC1-4.