Learning together: behavioral, computational, and neural mechanisms underlying social learning in adolescence

Westhoff, B.

5

Chapter 5

Increased ventromedial prefrontal cortex activity in adolescence benefits prosocial reinforcement learning

This chapter is published as:

Westhoff, B., Blankenstein, N. E., Schreuders, E., Crone, E. A., *&* van Duijvenvoorde, A. C. K (2021). Increased ventromedial prefrontal cortex activity in adolescence benefits prosocial reinforcement learning. *Developmental Cognitive Neuroscience*, 52, 101018. https://doi.org/10.1016/j.dcn.2021.101018

Abstract

Learning which of our behaviors benefit others contributes to forming social relationships. An important period for the development of (pro)social behavior is adolescence, which is characterized by transitions in social connections. It is, however, unknown how learning to benefit others develops across adolescence and what the underlying cognitive and neural mechanisms are. In this functional neuroimaging study, we assessed learning for self and others (i.e., prosocial learning) and the concurring neural tracking of prediction errors across adolescence (ages 9-21, N=74). Participants performed a two-choice probabilistic reinforcement learning task in which outcomes resulted in monetary consequences for themselves, an unknown other, or no one. Participants from all ages were able to learn for themselves and others, but learning for others showed a more protracted developmental trajectory. Prediction errors for self were observed in the ventral striatum and showed no age-related differences. However, prediction error coding for others showed an age-related increase in the ventromedial prefrontal cortex. These results reveal insights into the computational mechanisms of learning for others across adolescence, and highlight that learning for self and others show different age-related patterns.

Introduction

Adolescence is a developmental phase that is characterized by transitions in social connections, and moreover, a phase during which social cognitive skills are acquired and/or improved (Blakemore & Mills, 2014; Casey et al., 2008; Crone & Dahl, 2012; Sawyer et al., 2018). As social acceptance and approval from peers often result from displaying prosocial behaviors, for adolescents establishing their social network it is key that they learn to help or benefit others (Steinberg & Morris, 2001). That is, to be able to behave in a prosocial manner, individuals need to learn which actions would result in positive outcomes for others. This type of learning is also referred to as prosocial learning (Lockwood et al., 2016; Sul et al., 2015). Generally speaking, learning from actions and outcomes is an important part of cognitive development and continues to improve in adolescence (Bolenz et al., 2017; Nussenbaum & Hartley, 2019; Peters, Braams, et al., 2014; Peters et al., 2016). For adolescents, an especially salient environment that requires learning about the consequences of their actions is the interpersonal context (Blakemore & Mills, 2014; Nelson et al., 2005; Sawyer et al., 2018). Therefore, it is expected that especially prosocial learning shows improvements in adolescence. The goal of the current study was to unravel age-related differences in learning to benefit others using a prosocial learning context across adolescence.

The vast majority of recent neuroscientific studies investigating learning make use of formal reinforcement learning (RL) models. These models calculate individuals' prediction errors (PEs) – the difference between expected and actual outcomes - over the course of learning. These PEs drive learning via a learning rate, which quantifies to what extent these PEs affect subsequent actions. Consequently, RL models and the resulting PEs enable studies to examine the neural tracking of value-guided decision-making. Neuroscientific studies demonstrated that PE coding in a probabilistic reinforcement task context is associated with activation in the ventral striatum, as well as the medial prefrontal cortex (mPFC) (see for reviews e.g., (Cheong et al., 2017; Joiner et al., 2017; Lockwood & Klein-Flügge, 2020; Olsson et al., 2020; Ruff & Fehr, 2014). Developmental studies using RL models found that adolescents show similar neural tracking of PEs as adults when learning stimulus-outcome associations. However, the developmental patterns are inconsistent: some studies have reported elevated or lowered PE activity in the ventral striatum and connected structures in mid-adolescents relative to children and adults (Cohen et al., 2010; Davidow et al., 2016; Hauser et al., 2015; Jones et al., 2014), but this is not replicated in all studies (Christakou et al., 2013; van den Bos, Cohen, et al., 2012). Furthermore, age-related differences have been found in functional connectivity between the ventral striatum and mPFC, here referred to as ventromedial PFC (vmPFC), in relation to learning (van den Bos, Cohen, et al., 2012), suggesting that age-related improvements in learning are associated with stronger neural coupling between subcortical

and cortical brain regions (van Duijvenvoorde et al., 2016, 2019). Taken together, previous studies point to the ventral striatum and medial prefrontal cortex as important brain areas for learning in non-social environments.

Previous studies investigating the neurocomputational mechanisms of prosocial learning have investigated whether the same neural signaling occurs for PEs for others as for self. Recently, in adults, it was found that PE tracking for both learning for others as for self occurred in the ventral striatum (Lockwood et al., 2016). However, the subgenual anterior cingulate cortex (sgACC) specifically coded PE tracking for learning for others, and these prosocial learning signals were predicted by cognitive empathy. That is, more empathic people showed more activity in the sgACC when learning to benefit others. Cognitive empathy – the ability to understand the emotional states of others (Netten et al., 2015; Pouw et al., 2013) - shows pronounced changes in adolescent development and relates positively to prosocial behaviors such as trust and reciprocity (Dumontheil et al., 2010; Eisenberg et al., 1995; van de Groep et al., 2018). Therefore, we aimed to extend prior work by Lockwood and colleagues (2016) by investigating the neural tracking of PEs for others, and its relation with individual differences in cognitive empathy, in an adolescents sample with participants aged between 9 and 21 years.

In the current study, we adopted a prosocial learning task (Lockwood et al., 2016) in which participants could learn to obtain rewards for themselves, others, or no one. We administered this task to 74 adolescents between ages 9-21 years to examine age-related differences in learning for self and others, combined with functional neuroimaging (fMRI) for neural tracking of PEs. We use the term adolescence for this broad age range, based on definitions that mark adolescence from the onset of puberty to the age when one reaches independence from parents (i.e., approximately 9-24 years; e.g., Sawyer et al., 2018). Based on prior studies, we performed regions-of-interest analyses for the ventral striatum, sgACC, and vmPFC. We expected that adolescents, similar to adults, would show PE related neural activity when learning both for self and others in the ventral striatum (Lockwood et al., 2016), and in the sgACC and possibly vmPFC when learning for others more than when learning for self (Christopoulos & King-Casas, 2015; Lockwood et al., 2016). For learning for self, research has remained inconclusive whether this activity peaks in mid-adolescence (Cohen et al., 2010; Davidow et al., 2016) or shows no age-related differences (van den Bos, Cohen, et al., 2012). Therefore we explored linear as well as non-linear (quadratic) age effects. We predicted that sgACC and vmPFC activity for prosocial learning would increase with age, based on prior studies showing age-related improvements in social-cognitive perspective-taking (Dumontheil et al., 2010). Finally, consistent with (Lockwood et al., 2016), we expected that individual differences in cognitive empathy would relate to neural tracking of PEs for others.

Methods and Materials

Participants

A total of 76 participants between ages 9 and 21 took part in this study. Participants were recruited through schools and local advertisements, as well as from participation in a previous study. Two participants were excluded from analyses because they were either diagnosed with a psychiatric disorder at the time of testing (n = 1) or because the session was stopped early due to discomfort in the scanner (n = 1). We did not exclude participants based on task performance; there were no significant outliers in task performance (i.e., >3 SD) in any of the conditions. Four participants missed one run of the task, due to technical issues (n = 2), or discomfort in the scanner (n = 2). These four participants were maintained with the available data in all analyses. The final sample included 74 healthy participants (39 female, M_{age} = 15.64, SD_{age} = 4.18, range = 9.03 – 21.77 years, see Figure S1 for an overview of the number of participants across ages). The IQ scores, estimated with the Similarities and Block Design subtests of the WISC-III and WAIS-III, fell within the normal range (M_{IQ} = 110.24, SD_{IQ} = 10.37, range = 87.50 - 135.00), and did not correlate with age ($r(72)$ = -0.11, p = .353).

The local institutional review board approved this study (reference: NL56438.058.16). Adult participants and parents of minors provided written informed consent, and minors provided written assent. All anatomical scans were cleared by a radiologist and no abnormalities were reported. Participants were screened for MRI contraindications and psychiatric or neurological disorders, and had normal or corrected-to-normal vision.

Prosocial learning task

Participants played a two-choice probabilistic reinforcement learning task (prosocial learning task) in the MRI scanner (see Figure 1A). Participants were instructed to make a series of decisions between two pictures. One picture was associated with a high probability of winning 1 point, the other picture with a high probability of losing 1 point. The exact probabilities were 75% and 25% but were unknown to the participant. After the decision, participants were presented with the outcome to enable them to learn the reward contingencies.

The participants played the task in three different conditions: for themselves (Self), for an unknown other participant (Other), or for No One. The latter condition was added as a control condition based on Lockwood et al. (2016). Participants did not meet the other person, but were told that the other person was a peer also participating in the experiment who i) would not play the same game for them, ii) did not know who played for them (see Participant instructions in the Supplementary materials). Each block started with an instruction screen that indicated who would receive the outcomes (Self, Other, or No One) for 2000 ms. This was followed by the presentation of two stimuli for 2500 ms during which participants were required to select one of these. The stimuli were common objects, such as chairs, apples, and

shoes (see also Van Den Bos et al., 2009). If no response was given within the time frame, the text "Too late" appeared in the middle of the screen, and these trials were excluded from analyses.

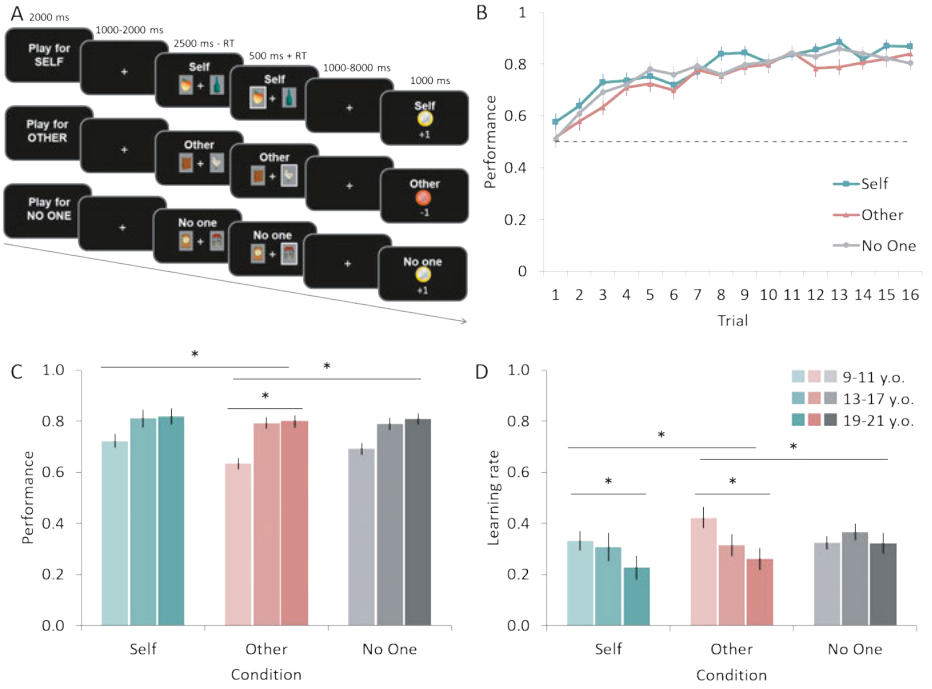

Figure 1. Prosocial learning task and behavioral data. **(A)** Participants played a two-choice probabilistic reinforcement learning task in which outcomes resulted in monetary consequences for themselves (Self condition), for an unknown other participant in the experiment who could not reciprocate (Other condition), or for No One. **(B)** Group-level performance across trials (learning curves) per condition, averaged across blocks. Performance represents the fraction of selecting the stimulus with a high reward contingency. The dashed line indicates performance at chance level (0.5). **(C)** Performance per condition per age cohort, averaged across the entire task. In all conditions, performance improved across trials, but an age-related increase was only observed when learning for others. Note that age is used as a continuous variable in all analyses but is visualized as age cohorts for illustrative purposes. The age-related increase was greater for the Other than for the Self and No One condition. **(D)** Learning rates per condition per age cohort. Age-related decreases in learning rates are only observed in the Self and Other condition. The age-related decrease in learning rate was greater in the Other compared to the Self and No One condition. Asterisks indicate significant effects. Error bars represent standard error of the mean (s.e.m.).

A selection frame around the chosen picture confirmed the response and remained visible for the duration of the interval and an additional 500 ms. A fixation screen (duration randomly jittered between 1000-2000 ms) preceded the outcome of their choice (+1 point or -1 point; 1000 ms). A randomly jittered fixation screen (1000-8000 ms) was shown after the outcome before the two pictures were presented again. The screen position of the stimulus (left or right) was counterbalanced across trials. Participants were instructed that the position of the stimulus did not matter, to encourage them to learn the reward contingencies regardless of stimulus position.

There were 144 trials in total, 48 for Self, 48 for Other, and 48 for No One, presented in three blocks of 16 trials. Each block began with a new pair of pictures. Participants completed three separate fMRI runs with a short break in between, each with one block of 16 trials per condition. The order of the conditions was counterbalanced across runs and between participants.

Participants were instructed that the total number of points in the Self condition was converted to money (each point valued €0.25), which they would get paid out on top of their flat participation rate (€20 for 9-11 y.o., €25 for 13-17 y.o., and €30 for 19-21 y.o.). The minimum of this extra amount of money was €1 to avoid null scores, and the maximum was €12. Additionally, participants were instructed that their choices in the Other condition were paid out to a participant entering the experiment after them. Consequently, participants received an additional fee from a participant before them in the experiment (minimum €1 and maximum €12), but only at the end of the experiment. Finally, it was instructed that choices in the No One condition had no financial consequences.

Cognitive empathy

To assess cognitive empathy, participants completed the Interpersonal Reactivity Index (IRI; Davis, (1983). This widely used self-report questionnaire consists of 4 subscales (Perspective-Taking and Fantasy as cognitive empathy subscales; and Personal Distress and Empathic Concern as affective empathy subscales) with 6 items each. To create a measure of cognitive empathy, two subscales were combined (Pulos et al., 2004): the Perspective-Taking subscale (e.g., "I sometimes try to understand my friends better by imagining how things look from their perspective", Cronbach's alpha = 0.710) and the Fantasy subscale (e.g., "I really get involved with the feelings of the characters in a novel", Cronbach's alpha = 0.786). All items can be answered on a five-point Likert scale ranging from (0) not true at all to (4) completely true, and higher scores indicate higher levels of empathy. Cognitive empathy scores increased across age (r = .309, p = .008, see Figure S8). One person did not fill in this questionnaire. This person was excluded from further analyses concerning measures of (cognitive) empathy. We

used a Dutch adolescent version for all ages in our study, with items adapted for the youngest ages in the study (Hawk et al., 2013).

Procedure

Participants were accustomed to the MRI environment using a mock scanner, and received instructions on the prosocial learning task in a quiet laboratory room. Instructions for the task were displayed on a screen and read out loud by an experimenter. Participants completed 8 practice trials in each condition. In the scanner, participants responded with their right hand using a button box. Head movements were restricted with foam padding. The fMRI scan was accompanied by a high-definition structural scan. Questionnaires were filled out at their home prior to the scanning session, via Qualtrics (www.qualtrics.com).

Computational modeling of behavioral data

Model fitting

We used MATLAB 2015b (The MathWorks Inc) for all model fitting and comparison. We modeled learning behavior in the Self, Other, and No One conditions separately, using a standard Rescorla-Wagner reinforcement learning (RL) model (similar to Lockwood et al., 2016) to obtain PEs and learning rates, which were subsequently used in behavioral and fMRI analyses. Simple RL models state that the expected value of a future action ($Q_{t+1}(i)$) should be a function of current expectations ($Q_t(i)$) and the difference between the actual reward that has been experienced on this trial (R_t). The learning rate α, bounded between 0 and 1, determines how much the value of the chosen stimulus is updated based on the new outcome. In particular, the learning rate parameter speeds up or slows down the acquisition and updating of associations. Optimal learning rates differ between contexts and reinforcement structures (Nussenbaum & Hartley, 2019).

$$Q_{t+1}(i) = Q_t(i) + \alpha * [\underbrace{R_t - Q_t(i)}_{\text{prediction error}}]$$

To select an action based on the computed values, we used a standard softmax choice function. For a given set of parameters, this equation allows us to compute the probability of the next choice being "i":

$$P_t(i) = \frac{e(\beta * Q_{i,t})}{\sum_j e(\beta * Q_{j,t})}$$

Beta (β) determines *how strongly* action probabilities are guided by their expected values (Q). Here, with larger β, actions are more deterministic and driven by expected values, resulting in selecting the option with the highest value. With lower β, actions are more random or exploratory. This parameter thus affects errors, where a decrease will lead to more random (i.e., less driven by expected values) choices. β did not differ between conditions, although with age, people were more strongly driven by expected values (see Figure S2 for the β across age cohorts for each condition).

We used the maximum a posteriori (MAP) approach (Daw, 2011) for fitting the RL model to participants' choices per condition. To facilitate stable estimation across subjects, we used weakly informative priors to regularize the estimated priors toward realistic ones. These weakly informative priors and estimation procedures were based on previous research (den Ouden et al., 2013), and included a Beta (1.2, 1.2) distribution for the estimated α (learning rate) parameter ($0 < \alpha < 1$) and a Gaussian distribution (0, 10) for the estimated β parameter ($-\infty \leq \beta \leq \infty$). Mean and confidence intervals for each of the fitted parameters across all subjects are displayed in Supplementary Table S1.

Model comparison

Based on previous developmental findings (e.g., van den Bos et al., 2012) we compared an alternative model with two learning parameters (i.e., separate learning rates for gains and losses) in order to benchmark the performance of the one-learning parameter model (i.e., one learning rate). Model comparisons revealed that the one-learning parameter model had a superior fit to the behavioral data for each condition, according to the Bayesian Information Criterion (BIC) (see Figure S3). This was the case in each condition for the majority of the participants (81.1% Self, 74.3% Other, 76.7% No One), in all age cohorts, see Figure S3. In none of the conditions, the BIC difference scores (Figure S3) were correlated with age (r_s, all p values > .14).

Simulations and parameter recovery

To assess whether computational model parameters could be successfully recovered, we simulated choice behavior for the range of learning rates and beta's that we encountered in our dataset. That is, we simulated a new participant dataset based on the α and β values from our participants as input parameters. This resulted in a simulated dataset with 74 participants. Parameter recovery, as indicated with correlations between simulated and recovered learning rates and beta values per conditions, is presented in Figure S4.

Behavioral analyses

To assess learning for Self, Other, and No One, and their developmental patterns in the prosocial learning task, we fitted logistic generalized linear mixed models (GLMMs) to decisions

(correct coded as 1, incorrect as 0) for each condition separately. These analyses were conducted in R version 4.0.1 (R Core Team, 2020), using the lme4 package (Bates et al., 2014). Our GLMMs included fixed effects of Age in years (linear and quadratic), Condition, Trial, and all interactions. Since no significant main or interaction effects of age-quadratic were observed in the choice data, this term was dropped in the final presented behavioral models for model parsimony. In all models, participant ID entered the regression as a random effect to handle the repeated nature of the data. Where applicable, Trial was additionally included as a random slope per subject. We performed post hoc tests using the *emmeans* package (Lenth et al., 2021), as well as tests per condition to delineate Age x Trial x Condition effects.

Next we examined the estimated learning rates per condition. These parameters indicate how people updated the value of stimuli based on outcomes for Self, Others, and No One. Since learning rates were not normally distributed, we used a robust linear mixed effects model (RLMM, rlmer function, robustlmm package (Koller, 2016) in R (see also Cutler et al., 2021), with Condition and Age linear as fixed main effects and interaction effects. We performed post-hoc tests per condition and pair-wise contrasts per Condition. In all GLMM and RLMM models, continuous independent variables were mean-centered and scaled, and categorical predictor variables were specified by a sum-to-zero contrast (e.g., sex: -1 = boy, 1 = girl). P-values for the GLMM were generated by using the Anova log-likelihood ratio tables from the afex package (Singmann et al., 2019). For the RLMM models, the Satterthwaite-approximated degrees of freedom generated by the lme4 model in combination with the output of the RLMM, was used to generate P-values.

Finally, we assessed whether cognitive empathy related to learning performance, learning rate, and PE activation when learning for others. We ran (partial) spearman correlational analyses with learning for self and cognitive empathy as predictors using the package 'ppcor' (Kim, 2015).

fMRI acquisition

For acquiring (functional) MRI data, we used a 3T Philips scanner (Philips Achieva TX) with a standard eight-channel whole-head coil. The learning task was projected on a screen that was viewed through a mirror on the head coil. Functional scans were acquired during three runs of 200 dynamics each, using T2* echo-planar imaging (EPI). The volumes covered the entire brain (repetition time (TR) = 2.2 s; echo time (TE) = 30 ms; sequential acquisition, 38 slices; voxel size 2.75 x 2.75 x 2.75 mm; field of view (FOV) = 220 (ap) x 220 (rl) x 114.68 (fh) mm). The first two volumes were discarded to allow for equilibration of T1 saturation effects. After the learning task, a high-resolution 3D T1 scan for anatomical reference was obtained (TR = 9.76 msec, TE = 4.95 msec, 140 slices, voxel size = 0.875 x 0.875 x 0.875 mm, FOV = 224 (ap) x 177 (rl) x 168 (fh) mm).

Preprocessing

Data were analyzed using SPM8 (Wellcome Department of Cognitive Neurology, London). Images were corrected for slice timing acquisition and rigid body motion. We spatially normalized functional volumes to T1 templates. Occasional framewise displacement >3mm occurred for 3 participants in 1-2 volumes. For those participants with frame-frame head motion >3mm, an extra regressor was included corresponding to each volume (n = 3, for maximum 2 volumes). All other participants did not exceed translational head movement more than 3mm in any of the scans (*Mean* = 0.65mm, *SD* = 0.059mm). The normalization algorithm used a 12 parameter affine transform with a nonlinear transformation involving cosine basis function, and resampled the volumes to 3 mm^3 voxels. Templates were based on MNI305 stereotaxic space. The functional volumes were spatially smoothed using a 6 mm full width at half maximum (FWHM) isotropic Gaussian kernel.

General linear model

We used the general linear model (GLM) in SPM8 to perform statistical analyses on individual subjects' fMRI data. The fMRI time series were modeled as a series of two events: the decision phase (Expected Value, EV) and the outcome phase (PE), convolved with a canonical hemodynamic response function (HRF). The onset of the choice (EV), and the onset of the outcome (PE) were both modeled with zero duration. Each of these regressors was associated with a parametric modulator taken from the computational model. At the time a stimulus was selected (decision phase) this was the chosen expected value, and at the time of the outcome, the PE. The PEs were estimated using each subject's own alpha and beta from each condition. Trials on which participants did not respond were modeled separately as a regressor of no interest. Six motion parameters, and -if applicable- motion censoring regressors were included as nuisance regressors. We used the MarsBaR toolbox (Brett, Anton, Valabregue, & Poline, 2002; http://marsbar.sourceforge.net) to visualize the patterns of activation, in clusters identified in the whole-brain results. Coordinates of local maxima are reported in MNI space. Our main hypotheses centered on PE coding. For completeness, effects of EV at choice onset are included in Supplementary Table S3. In addition, uncorrected T-maps of EV and PE effects are uploaded on Neurovault (https://neurovault.org/collections/EOTSVZYT/). For condition effects, we examined contrasts of Self versus Other in concordance with Lockwood et al. (2016). Contrasts were obtained from a flexible factorial design with three levels (Self PE, Other PE, No One PE). Effects and conclusion remained the same when testing Self PE > Other PE + No One PE, and Other PE > Self PE + No One PE. In Supplementary Table S4 we include all contrasts between conditions within our ROIs. Whole-brain effects for main effects and between conditions are included in Supplementary Tables S2 and S5, respectively. Age effects (linear and quadratic) were tested in follow-up regressions.

ROI selection and fMRI analyses

The a priori regions of interest (ROI) in which we test our main hypotheses were defined anatomically and based on previous research on (prosocial) learning and feedback processing (Lockwood et al., 2016; van den Bos, Cohen, et al., 2012; van Duijvenvoorde et al., 2014). In concordance with previous studies, masks were taken from an appropriate atlas. That is, the bilateral ventral striatum and vmPFC were determined by an anatomical mask from the Harvard-Oxford Atlas (van Duijvenvoorde et al., 2014; van den Bos et al., 2012; Braams et al., 2015; Peters *&* Crone, 2017), and the sgACC was defined as Brodmann areas (BA) 25 and s24 (Lockwood et al., 2016). The sgACC region and the ventral striatum are anatomically adjacent and partly overlapping (see Figure S5), but significant peak activations in either ROI were not observed in these overlapping voxels. Coordinates for local maxima are reported in MNI space. Effects in our ROIs are reported at $p < .05$ FWE-small volume corrected (SVC). Predictions were tested while correcting for multiple comparisons (3 ROIs) by limiting the false discovery rate (FDR; Benjamini *&* Hochberg, 1995); all reported tests survived this correction. Explorative whole-brain analyses are reported in Supplementary Tables S2 and S5, and Figure S6).

Results

Developmental differences in learning to obtain rewards for Self, Others, or No One

Results showed that, at the group level, participants were able to learn for Self, Other, and No One, as they performed above chance level in all conditions (0.5; *t* values > 13.0, all *p*s < .001, *df* = 73; Figure 1B). Using a generalized linear mixed model (GLMM) on participants' choice behavior over trials, we assessed age-related differences in performance when learning for Self, Other, and No One. Performance in the learning task improved linearly with age (main effect of Age linear, $p = .001$). Moreover, we observed that age-related differences in learning performance differed per condition (Age x Condition interaction, $p = .005$). Post-hoc analyses revealed that the age-related improvement in performance was larger when learning for Other than when learning for Self ($p = .009$) and when learning for No One ($p = .02$). The age-related improvements were similar for learning for Self and No One ($p = .92$). Similarly, we also observed age-related differences in learning curves across trials, which differed per condition (Age x Condition x Trial interaction, $p = .007$). Specifically, younger children learned more slowly (i.e., flatter learning curves) across trials when learning for others, but this age effect on trial was not observed for the Self and No One condition (Age linear x Trial, for Other condition, $p < .001$; Self and No One conditions: *p*s > 0.2; see Figure 1C and Figure S7). Together, these findings suggest that across adolescence prosocial learning shows a more protracted improvement than when learning for Self or No One.

Next, we examined participants' learning rates to assess how they updated the value of stimuli on the basis of outcomes for Self, Others, and No One. That is, higher learning rates indicate that people adjusted behavior quickly towards recent feedback, whereas lower learning rates indicate a slower pace in updating in which outcomes across multiple trials are integrated. Using a robust linear mixed effects model, we assessed effects of Condition and Age (linear) in learning rates (Figure 1D). We observed that learning rates for Self were lower than learning rates for Other ([Self vs Other], $b = 0.02$, $p < .001$) and for No One ([Self vs No One], $b = -0.03$, $p < .001$). Learning rates for Other and for No One did not differ ([Other vs No One], $p = .911$). Moreover, we observed that learning rates decreased linearly with age (main effect of Age linear, $b = -0.04$, $p = .023$), an effect that also differed across conditions. Specifically, learning rates decreased across age in the Other and Self condition, but more strongly across age for Other than for Self ([Other-Self]*Age, $b = -0.02$, $p < .001$) and for Other than for No one ([Other-No One]*Age, $b = 0.004$, $p = .004$). Learning rates also decreased more strongly across age for Self than for No One ([Self-No One *Age, $b = .019$, $p < .001$). Learning rates did not differ across age in the No One condition ($p = .08$). Together, these findings show that for both learning for Self and Others, younger participants responded more to recent feedback, whereas older participants integrated feedback more over trials. Moreover, this age-related change was most pronounced in the Other compared to the Self and No One condition.

Identifying common and distinct coding of prediction errors for Self and Others

To formally investigate the brain regions that were responding to PEs for Self, Others, and No One, we conducted a conjunction analysis to explore whether there were regions that commonly code PEs across all conditions. Common activation for PEs regardless of the beneficiary was observed in the vmPFC (MNI coordinates [$x = -9$, $y = 44$, $z = -11$], $Z = 5.33$, $k = 136$, $p < .001$, SVC-FWE), ventral striatum ([$x = -9$, $y = 11$, $z = -11$], $Z = 5.05$, $k = 23$, $p < .001$, SVC-FWE, and [$x = 12$, $y = 14$, $z = -8$], $Z = 4.43$, $k = 18$, $p < .001$, SVC-FWE), and sgACC ([$x = -6$, $y = 14$, $z = -8$], $Z = 5.47$, $k = 32$, $p < .001$; and Self [$x = 6$, $y = 17$, $z = -8$], $Z = 4.60$, $k = 21$, $p = .001$; and [$x = 9$, $y = 8$, $z = -14$], $Z = 3.67$, $k = 2$, $p = .029$, SVC-FWE) (see Figure 2). These findings show that all regions of interest were involved in PE coding, in each condition.

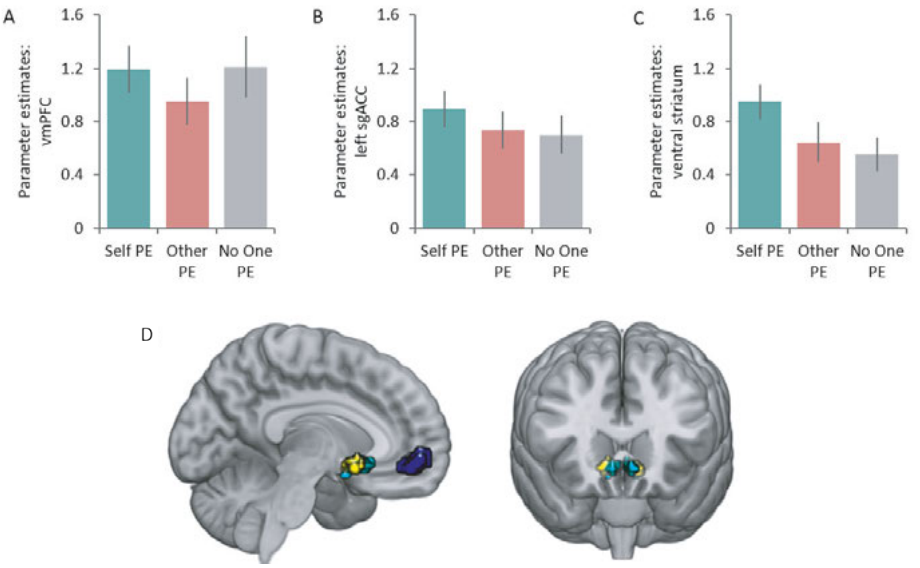

Figure 2. Common prediction error (PE) coding in three regions of interest. Shown are the responses to prediction errors for Self, Other, and No One in **(A)** the vmPFC, **(B)** left sgACC, and **(C)** ventral striatum. **(D)** Significant clusters of activation in the vmPFC (blue), sgACC (cyan), and ventral striatum (yellow). All images displayed at $p < .05$ FWE-SVC.

Next, we examined which brain regions responded more to PEs for Self than for Other by contrasting the Self condition against the Other condition (see Supplementary Table S4 for contrasts including the No One condition). The left ventral striatum was the only region to respond more strongly to PEs for Self ([x = 12, y = 11, z = -11], Z = 4.37, k = 9, $p < .001$, SVC-FWE; Figure 3). When examining effects of age we observed no linear or quadratic age-related differences in self-related PE coding. These findings indicate that the ventral striatum responds more to PEs for Self than for Others, and this effect did not differ across age.

We next identified regions that corresponded to PEs for others exclusively by contrasting the Other condition against the Self condition. No voxels in our ROIs responded more strongly to prosocial PEs than Self PEs. When adding age (linear and quadratic) to the model to examine whether age-differences were related to prosocial PE coding, we observed that the vmPFC increasingly responded to prosocial PEs with age ([x = -15, y = 50, z = 8], Z = 4.95, k = 45, p = .004, SVC-FWE; see Figure 4). No effects of quadratic age were observed. This shows that the vmPFC is increasingly involved in prosocial PE coding across adolescence.

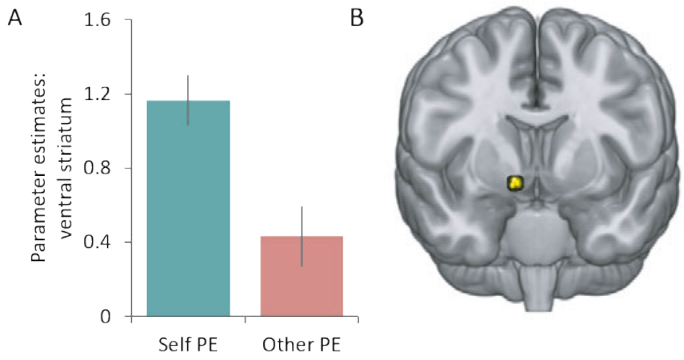

Figure 3. Ventral striatum response to prediction errors for Self versus Other. **(A)** Left ventral striatum [x=12, y=11, z=-11] response for Self PE and Other PE. **(B)** Overlay of the response for Self PE > Other PE in the left ventral striatum. All images displayed at $p < .05$ FWE-SVC.

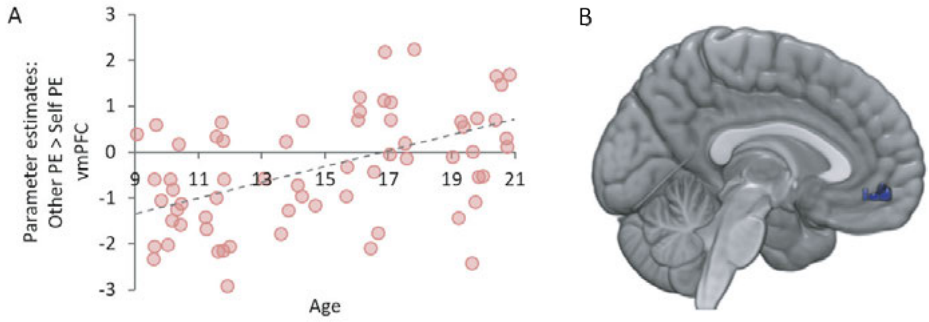

Figure 4. Linear age effects in responses to Other PE > Self PE in the vmPFC. **(A)** scatterplot showing the relation between age and activation in the vmPFC for Other PE > Self PE. Scatterplot is only presented for visualization. **(B)** Overlay of the response for Other PE > Self PE in the vmPFC [-15, 41, -11]. All images displayed at $p < .05$ FWE-SVC.

Links between cognitive empathy and learning for Others

Finally, we examined the link between cognitive empathy and prosocial learning. First, we assessed whether cognitive empathy related to performance for Other, while controlling for performance for Self. We observed that individuals with higher empathy ratings, showed better prosocial learning ($r_s = .30$, $p = .01$). Subsequently, we assessed whether cognitive empathy related to learning rate in the Other condition (controlled for learning rate in the Self condition). Results showed that individuals with higher empathy ratings had lower learning rates when learning for Others (cognitive empathy, $r_s = -0.26$, $p = .027$, see Figure 5B). To-

gether, these findings indicate that individuals with more empathy show better learning performance, and integrate information more over trials when learning to benefit others. Finally, we assessed the relation between cognitive empathy and the prosocial PE coding in the vmPFC. For this purpose, we extracted the values of the Other PE > Self PE contrast in vmPFC that showed age-related change (see Figure 4). Results showed that greater Other vs Self-related PE activation in the vmPFC related to higher empathy scores (cognitive empathy, r_s = .31, p = .007).

To examine whether age-related differences in empathy or prosocial learning may influence these relations, we additionally included age in the partial correlation analysis. When additionally controlling for age, the relation between empathy and learning for others remained significant (p = .029), the relationship between empathy and learning rate became trend-level (p = .06), and the relation between empathy and prosocial PE coding was no longer significant (p = .15).

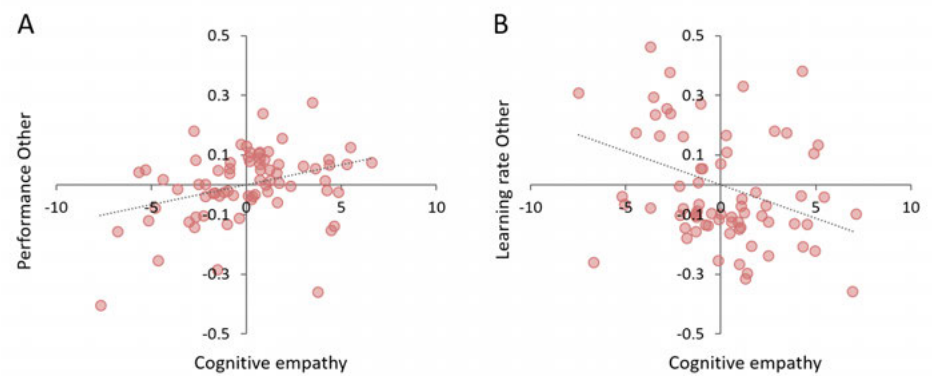

Figure 5. Relation of cognitive empathy with performance for Others and learning rate for Others. **(A)** Partial correlation plot showing that individuals with more cognitive empathy perform better for Others (controlled for performance for Self). **(B)** Partial correlation plot showing that individuals with more cognitive empathy have lower learning rates when learning for Others (controlled for learning rate for Self).

Discussion

This study examined the developmental trajectories of prosocial learning and self-related learning in an adolescent sample spanning ages 9-21 years. We examined the underlying mechanisms in this developmental sample by assessing the neural tracking of PEs during learning for self and others, and how individual differences in cognitive empathy relate to

prosocial learning performance. To this end, participants played a two-choice probabilistic reinforcement learning task in which outcomes resulted in monetary consequences for themselves (Self) or an unknown other (Other; prosocial). Our results show improvements in learning for self and others, but the developmental trajectory of prosocial learning is more protracted compared to learning for self. PEs for self were related to activation in the left ventral striatum, which did not show age-related differences. On the other hand, vmPFC-related PE activation during prosocial learning increased with age, and related to individual differences in cognitive empathy. Together, these findings highlight that learning for self and others show different age-related patterns.

The main goal of this study was to examine age-related differences in prosocial learning. Behaviorally, we observed that it is not until mid-adolescence that participants learn similarly for themselves and others. These findings may suggest a self-bias that is stronger in younger ages (van der Aar et al., 2018), and that the motivation to learn for self and others increases with age. Neurally, we observe that a reward-related network including the ventral striatum, sgACC, and vmPFC respond significantly to PEs when learning for Self, Other, and No One. This conjunction presented the starting point for our interest in testing condition-specific learning effects. Contrary to Lockwood et al. (2016), who observed similar PE neural tracking values in the ventral striatum for learning for Self and Others in adults, we observed that PE neural tracking was stronger in the ventral striatum for Self than for Others. Recent reviews, however, suggest that the striatum is related to a range of computations that take place during social learning that could reflect both self-related and other-related learning (Joiner et al., 2017), or the difference between winning for self and others (Báez-Mendoza & Schultz, 2013). Therefore, one explanation for our findings could be related to the possible stronger self-focus or the greater focus on social comparisons reflected in the ventral striatum.

Learning for Others, compared to learning for Self, was associated with stronger activation in the vmPFC with age. Previous prosocial reinforcement learning studies have suggested that the vmPFC is also responsive to processing of self-related expected values (Sul et al., 2015), self-representation (Sui & Humphreys, 2017), or does not differentiate between self and other-related PEs (Lockwood et al., 2016). On the other hand, the vmPFC is suggested to respond to prosocial rewards in adults (Christopoulos & King-Casas, 2015), to others' outcome PEs (Burke et al., 2010), and to simulated others' reward PEs (Suzuki et al., 2012). Our findings extend these prior studies by showing that the ventral striatum and vmPFC code PEs both for self and others (see also Joiner et al., 2017). In this first developmental sample investigating prosocial learning, we observe a specificity for Self PEs in the ventral striatum and an increased specificity for prosocial PE coding in the vmPFC, in which across age prosocial (compared to self-related) PE elicit more activation. Alternatively, the pattern of age-related differences we observed for Other and Self-learning in the vmPFC may also support the perspective of a decreasing self-focus with age. For instance, previous work on self-concept development

highlights that perspectives of others and self become more merged across development (van der Cruijsen et al., 2019). However, longitudinal studies are more powerful and essential for examining the true developmental trajectories of prosocial learning.

Besides the age-related differences in other-related learning, we observed that consistent with previous findings (Lockwood et al., 2016), individual differences in cognitive empathy were related to prosocial learning. Individuals with higher levels of empathy performed better for Others, integrated outcomes more over time (i.e., lower learning rates), and their vmPFC showed greater activation during prosocial PE coding. However, relations on cognitive empathy and prosocial PE coding in the brain were not robustly observed when controlling for age. This may indicate that it is hard to disentangle whether empathy or age drives prosocial PE coding. Also, age-related differences in brain activity during prosocial PE tracking may be explained by other social cognitive mechanisms than empathy. For instance, although there was no reciprocity or competition, participants may have been influenced by social inequality preferences, such as disliking to getting more (i.e., advantageous inequality aversion), or less (i.e., disadvantageous inequality aversion) than the other participant (Dawes et al., 2007; Fehr *&* Schmidt, 1999; Meuwese et al., 2015; Westhoff et al., 2020). Future studies could more explicitly assess several social-cognitive skills, strategies, and motivations along with a prosocial learning task to examine what behavioral mechanisms rely most on adolescents' prosocial learning.

Prior developmental studies on general reinforcement learning remained inconclusive about whether age-related differences were observed in PE neural tracking in the ventral striatum (Christakou et al., 2013; Cohen et al., 2010; Hauser et al., 2015; van den Bos, Cohen, et al., 2012). Here, age-related differences in PE coding for Self were not observed in the ventral striatum. In contrast to other studies (Cohen et al., 2010; Peters *&* Crone, 2017) we also did not find any quadratic age effects in learning or PE coding. This is possibly due to our narrower age range (9-21 y.o. instead of 8-30 y.o.), as another developmental study on learning also has not observed age-related changes in ventral striatum activity in a similar age range (van den Bos, Cohen, et al., 2012). Indeed, a recent review recommended using samples with wider age ranges, including children and adults, when examining quadratic age effects across adolescence (Li, 2017). It should be noted, however, that although we did not find age-effects in the ventral striatum, the behavioral learning performance for Self showed linear improvements with age. This could also indicate that other mechanisms than simple PE coding may be related to behavioral learning improvement over time within the current age range. For example, a prior study in young adults indicated that besides well-known model-free learning, another more sophisticated and flexible learning system is model-based learning. These two distinct computational strategies use different error signals which are computed in partially distinct brain areas (Gläscher et al., 2010). Moreover, it has been found that people may use different learning strategies, which show different neural activation patterns (Peters, Koolschijn, et al.,

2014). Future studies are needed to assess whether age-related improvements in learning performance may be more strongly related to strategic learning differences.

We observed that, overall, learning rates decreased with age, and lower learning rates were related to better performance. These findings indicate that, with age, adolescents increasingly integrate information across trials, which was beneficial to their prosocial learning performance. Intriguingly, a recent aging study with a similar prosocial learning task observed that better learning performance was related to *higher* learning rates instead (Cutler et al., 2021; Lockwood et al., 2016). Besides the included age range in this study, a few differences in modeling and task structure may underlie this deviance. First, we allowed a wide range of beta-parameters. Since beta-parameters showed consistent age-related declines (see Supplementary Figure 2), and also relate to performance (see Supplementary materials) this may have influenced our learning rate estimations. Second, the task structure shows differences in reinforcement structure. Most profoundly we included gains and losses compared to gain and no-gains in previous prosocial learning studies. Possibly, losses may influence the updating of values across trials differently, although we did not find evidence that gains and losses were weighted differently in learning across development. Future studies should further examine the influence of reinforcement structures on observed age-related differences in reinforcement learning.

The current study had several limitations that can be addressed in future research. First, prosocial learning was restricted to unknown others, and participants did not meet these others. Although we circumvented the potential effects of reputational concerns, it may have been more salient to include a confederate, as used in previous studies on prosocial learning in which participants played for a stranger who they met prior to the experimental task (Lockwood et al., 2016; Sul et al., 2015). Second, it would be interesting if future research would extend the prosocial learning task to other beneficiaries. Previous studies have shown that prosocial behaviors and their neural correlates in adolescence strongly depend on the beneficiary (e.g., (Brandner et al., 2020; Schreuders et al., 2018; van de Groep et al., 2020; Westhoff et al., 2020). Future studies should further examine whether such differences between beneficiaries are also visible in prosocial *learning* and whether this affects the concurrent neural tracking of PEs. Third, the neural results for the No One condition showed an intermediate pattern between learning for Self and Others, which is difficult to interpret. Behavioral analyses showed that participants generally performed well in this condition (i.e., not significantly different from learning for Self), even though no monetary reinforcers were given depending on task performance. Although including the No One conditions in our contrasts of interest did not alter our main findings, this condition was possibly interpreted by participants in different ways, in which some participants were internally motivated to perform well (e.g., (Satterthwaite et al., 2012). Finally, in line with previous research we used a model with separate learning rates per condition (Lockwood et al., 2016). Using this established model,

our results also revealed expected differences in learning rate between conditions. However, other studies also included comparison testing whether different learning rates or beta's are needed across different conditions (Cutler et al., 2021). Future studies may expand on these recent modeling procedure in prosocial learning in developmental and adult populations.

In conclusion, we found that prosocial learning showed age-related improvements across adolescence, suggesting a developmental shift from self-focus in early adolescence to self and other-focus in late adolescence and early adulthood (Crone & Fuligni, 2020). This developmental improvement was associated with stronger recruitment of the vmPFC for others compared to self. This study has implications for learning in social settings, such as educational contexts (Altikulaç et al., 2019), as well as for how children develop prosocial values when learning for unknown others. This study provides the first building blocks to understand age-related differences in how adolescents learn to benefit others.

Supplementary materials

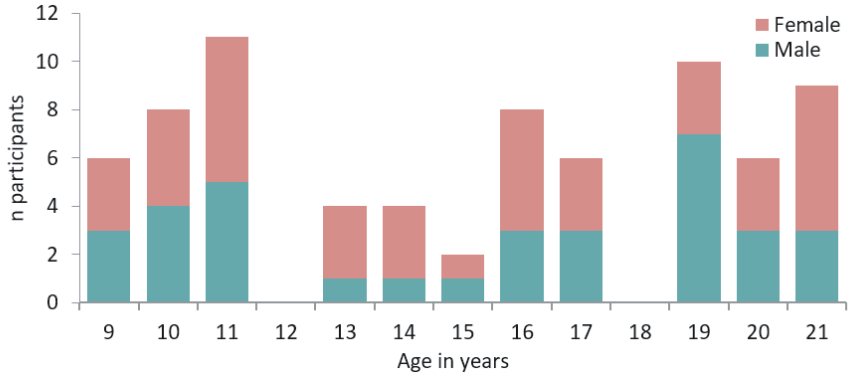

Figure S1. Number of participants across age per sex. In total, 74 participants were included (39 female, 35 male).

Beta parameter

The Beta parameter was examined as an index to what extent participants followed expected value in their choice behavior, and is also considered a parameter of decision noise. Higher values represent less decision noise here. Using a robust linear mixed effects model, we assessed effects of Condition and Age (linear) in beta parameters. We observed that with increasing age, decision noise decreased linearly (main effect of Age, $B = 2.9$, $p = .007$), indicating participants follow expected value more closely. Particularly beta parameters increased more strongly across age for Other than for Self (Other-Self; $B = 0.47$, $p < .001$), and did not differ significantly between No One and Other (No One – Other $p = .19$) and between No One and Self (No One – Self; $p = .19$)

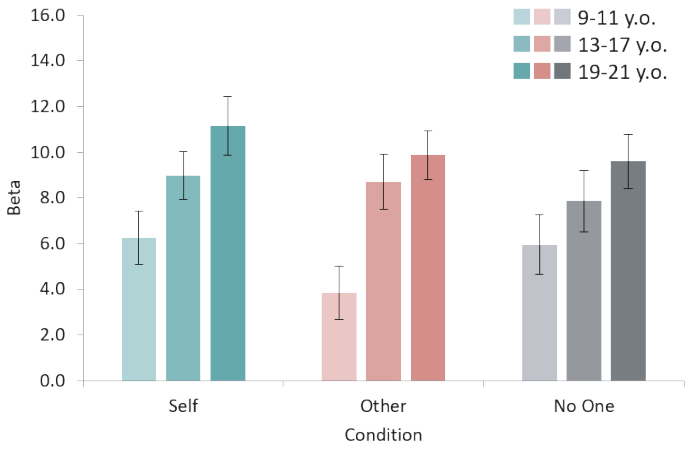

Figure S2. Beta parameter per condition per age cohort. Age is used as a continuous variable in all analyses, but is visualized as age cohorts for illustrative purposes and interpretability.

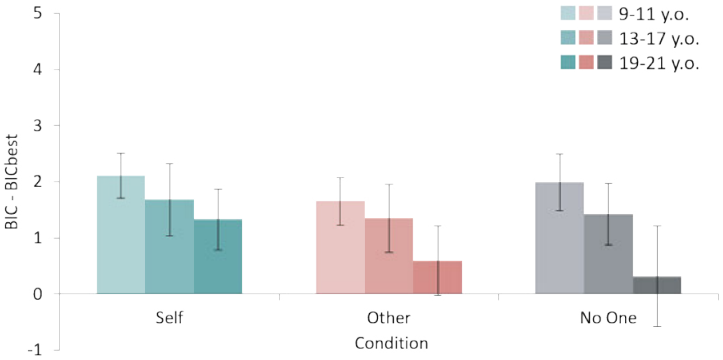

Figure S3. BIC values per condition per age cohort. Bars show BIC differences of the two-learning rate model (gain and loss) with the best model (one learning rate). Values on the y-axis indicate the difference between fit values (BIC values) for the two-learning rate model (gains and loss) and fit values for the one-learning rate model (the best model). BIC values were calculated per participant, and are shown separately per age cohort and per condition. For all age cohorts and all conditions a one-learning rate model is the best-fitting model. Lower bars indicate that the model fit of the one-learning rate model and two-learning rate model are more similar. For each condition, these BIC difference scores are not predicted by age (all *Ps* > .14). Error bars represent standard error of the mean.

Relations between performance, learning rates, and betas

We tested non-parametric correlations between performance, learning rates, and betas per condition. These show that lower learning rates in the Other condition are related to better performance for Other ($r_s(74) = -.38$, $p = .001$). Similarly, lower learning rates in the No One condition are related to better learning for No One ($r_s(74) = -.34$, $p = .003$, but learning rates in the Self condition are not significantly related to learning for Self ($r_s(74) = -.22$, $p = .056$). In addition, higher betas were strongly related to better performance in all conditions (Self, $r_s(74) = .91$, $p < .001$; Other $r_s(74) = .94$, $p < .001$; No One $r_s(74) = .89$, $p < .001$. Also, in all conditions, lower learning rates are related to higher betas (Self, $r_s(74) = -.37$, $p < .001$; Other, $r_s(74) = -.54$, $p < .001$; No One, $r_s(74) = -.48$, $p < .001$).

Figure S4. Learning rate and Beta parameter recovery. The correlation matrices represent the correlations between simulated and recovered **(A)** learning rates, and **(B)** beta values. Stronger colors show higher values and high values on the diagonal show parameters can be recovered.

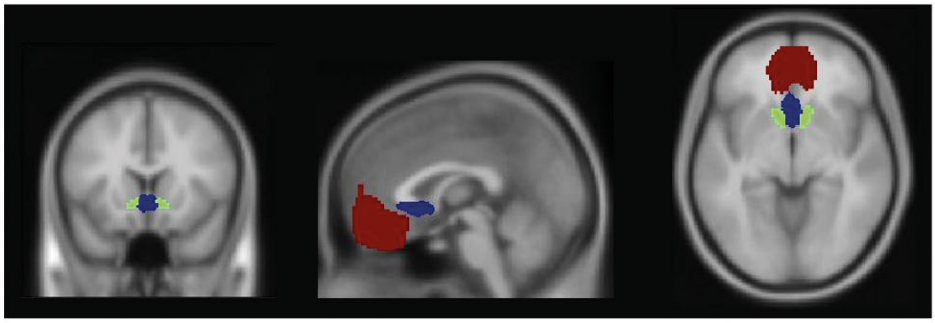

Figure S5. Regions of interest. Ventromedial prefrontal cortex (vmPFC; red), subgenual anterior cingulate cortex (sgACC; blue), and the ventral striatum (green).

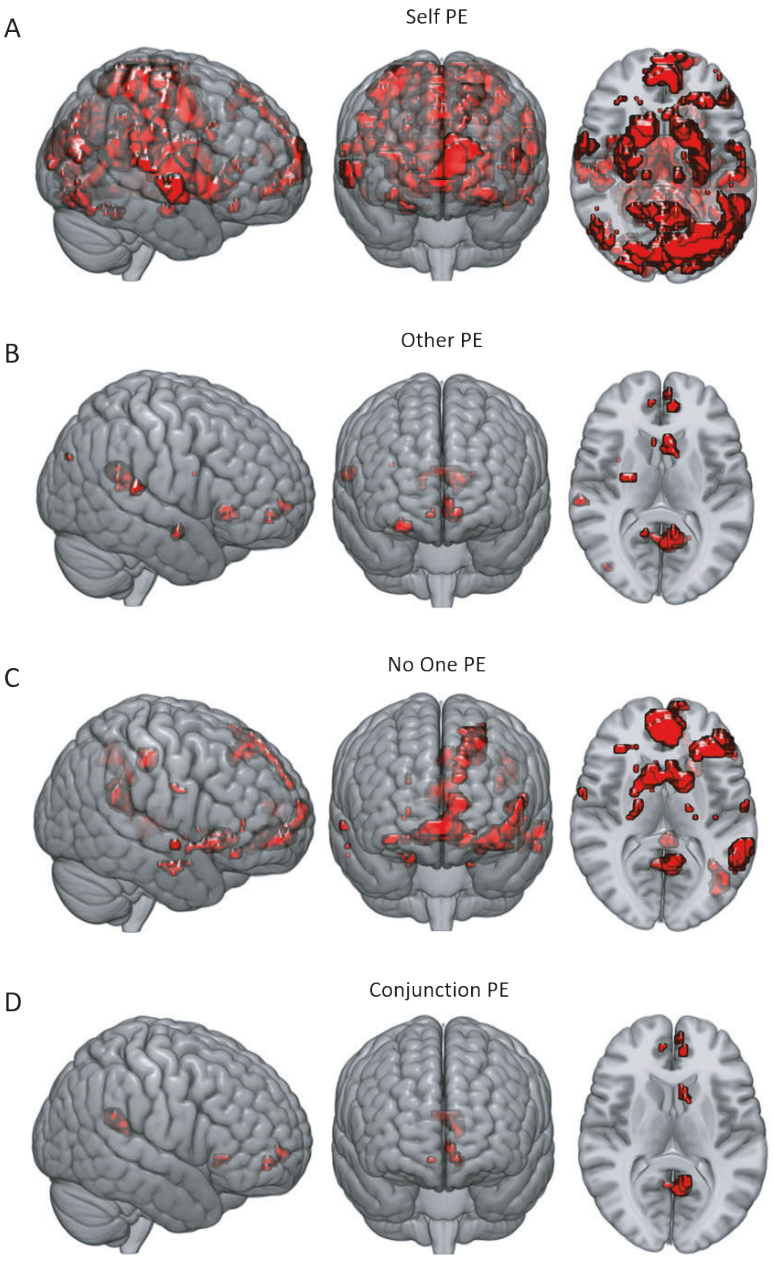

Figure S6. Whole brain responses for (A) Self PE, (B) Other PE, (C) No One PE, and (D) conjunction (common PE coding in all three conditions). All images displayed at $p < .05$ FWE, voxel level corrected.

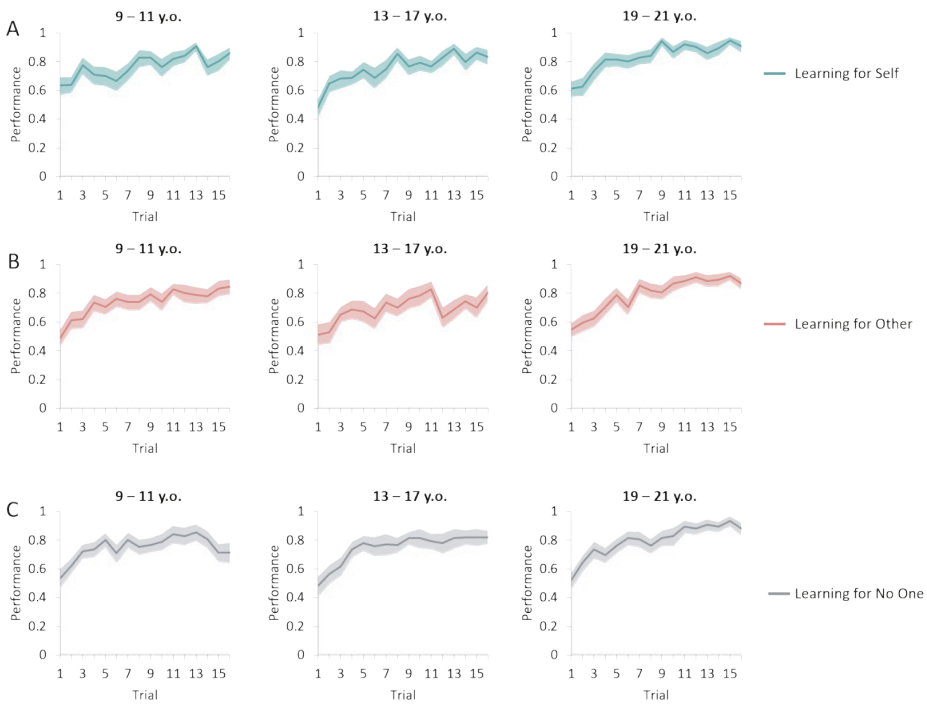

Figure S7. Learning across trials for (A) Self, (B) Others, and (C) No One, per age cohort. Note that for all analyses including age, we used age as a continuous variable. However, figures represent age per age cohorts instead for illustrative purposes and interpretability.

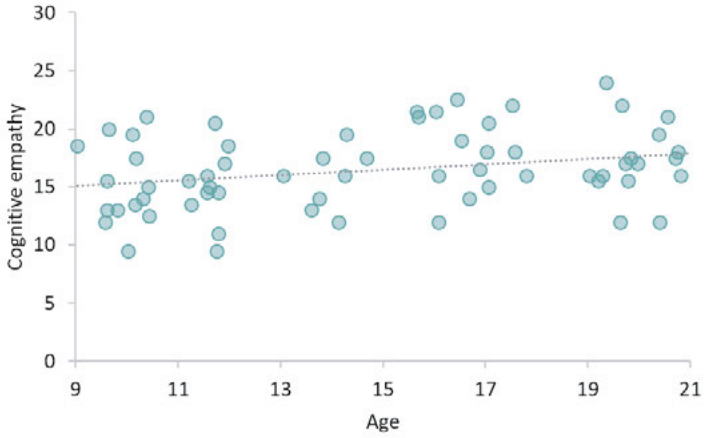

Figure S8. Cognitive empathy across age.

Table S1. Mean model parameters with prior distributions (M, SD), constraint, and 95% confidence intervals around the mean.

Model	Parameter	Prior	Constraint	Mean	95% CI
RL Self	α	β (1.2,1.2)	$0 < \alpha < 1$	0.28	.24 – .32
	β	Gaussian (0,10)	$-\infty \leq \beta \leq \infty$	8.81	7.37 – 10.26
RL Other	α	β (1.2,1.2)	$0 < \alpha < 1$	0.34	.29 – .39
	β	Gaussian (0,10)	$-\infty \leq \beta \leq \infty$	7.43	5.92 – 8.94
RL No One	α	β (1.2,1.2)	$0 < \alpha < 1$	0.34	.29 – .39
	β	Gaussian (0,10)	$-\infty \leq \beta \leq \infty$	7.94	6.52 – 9.37

Table S2. Main effects of whole brain prediction error responses per condition, and common prediction error coding (conjunction).

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	*t*	*z*
Main effect Self PE						
L Precental gyrus	-27	-25	61	6235	9.68	Inf
L precuneus	-3	-58	13		9.65	Inf
L Postcentral gyrus	-30	-34	64		9.03	Inf
L Putamen	-30	-13	4	481	9.52	Inf
L Putamen	-15	8	-11		8.20	7.65
R Caudate	12	11	-11	525	9.01	Inf
R Thalamus	30	-16	7		8.22	7.65
R Putamen	27	-7	10		8.04	7.52
L Superior frontal gyrus, medial	-9	65	19	515	7.25	6.85
L Superior frontal gyrus, medial	-12	58	7		6.93	6.58
L Superior frontal gyrus, medial	0	59	1		6.85	6.52
L Rolandic operculum	-48	-28	19	271	6.95	6.60
L Superior temporal gyrus	-60	-31	19		6.71	6.39
L Supramarginal gyrus	-60	-22	16		6.57	6.27
L Middle frontal gyrus, orbital part	-30	35	-14	50	6.73	6.41
L Thalamus	-12	-22	4	21	6.57	6.27
R Superior temporal gyrus	63	-1	-5	83	6.52	6.23
R Rolandic operculum	63	5	1		6.06	5.82
R Superior temporal gyrus	66	-7	4		5.29	5.13
L Inferior frontal gyrus,triangular part	-48	32	13	60	6.47	6.19
L Rolandic operculum	-54	-4	4	34	6.04	5.81
R Inferior temporal gyrus	51	-67	-11	46	5.76	5.55

Table S2. *Continued*

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	*t*	*z*
R Inferior temporal gyrus	51	-58	-20		5.41	5.23
R Inferior occipital gyrus	42	-76	-17		5.35	5.19
R Cerebellum	21	-52	-23	20	5.60	5.41
Main effect Other PE						
L Precuneus	-6	-61	13	103	6.54	6.25
L Calcarine fissure & surrounding cortex	-12	-55	10		6.19	5.94
L Olfactory cortex	-6	20	-11	32	6.29	6.02
R Hippocampus	30	-7	-20	17	5.66	5.46
R Superior temporal gyrus	63	-28	16	17	5.59	5.40
L Middle frontal gyrus, orbital part	-9	44	-11	12	5.58	5.39
Main effect No One PE						
L Middle frontal gyrus, orbital part	-3	50	-11	246	9.13	Inf
L Superior frontal gyrus, medial	-6	62	1		5.92	5.69
L Superior frontal gyrus, medial	-9	55	13		5.77	5.56
R Caudate	12	8	-11	173	7.64	7.19
L Olfactory cortex	-15	11	-14		6.72	6.40
L Olfactory cortex	-5	20	-11		6.19	5.93
L Inferior frontal gyrus, orbital part	-36	35	-14	182	7.60	7.15
L Middle frontal gyrus, orbital part	-24	32	-17		6.30	6.04
L Inferior frontal gyrus, triangular part	-45	32	7		5.99	5.75
L Middle temporal gyrus	-60	-43	-8	134	7.21	6.82
L Inferior temporal gyrus	-54	-52	-17		6.39	6.09
L Precuneus	-6	-55	16	120	7.16	6.78
L Calcarine fissure & surrounding cortex	-12	-52	7		5.88	5.66
L Median cingulate and paracingulate gyri	-3	-37	40	50	6.57	6.27
L Middle frontal gyrus	-24	32	49	172	6.56	6.26
L Middle frontal gyrus	-24	20	46		5.99	5.75
Superior frontal gyrus, medial	-12	59	28		5.64	5.45
L Angular gyrus	-42	-67	34	122	6.14	5.89
R Parahippocampal gyrus	18	-10	-26	16	5.63	5.43
R Parahippocampal gyrus	24	-19	-23		5.20	5.04
Conjunction						
L Precuneus	-6	-58	16	61	6.54	6.24
L Caudate	-6	14	-8	12	5.67	5.49

For all regions, FWE $p < .05$ voxel-level whole-brain corrected, and presented here with $k > 10$. PE = Prediction error; L = Left; R = Right; k =cluster extent. Names of the brain regions derived from the Automated Anatomical Labeling (AAL) atlas.

Table S3. Main effects of whole brain expected value responses per condition, and common expected value coding (conjunction).

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	*t*	*z*
Main effect Self EV						
Precuneus	-15	-64	19	14	5.63	5.44
Main effect Other EV						
L Precuneus	-9	-61	19	53	6.01	5.77
R Middle frontal gyrus, orbital part	3	56	-5	23	5.51	5.32
Main effect No One EV						
L Middle temporal gyrus	-51	-13	-8	11	5.36	5.19
Conjunction*						
L Precuneus	-12	-58	13	140	4.28	4.18
L Precuneus	-12	-58	22		3.99	3.91
R Precuneus	6	-58	19		4.13	4.04

For all regions, FWE $p < .05$ voxel-level whole-brain corrected, and presented here with $k > 10$. PE = Prediction error; L = Left; R = Right; k =cluster extent. Names of the brain regions derived from the Automated Anatomical Labeling (AAL) atlas. *threshold $p < .001$

Table S4. Comparison of responses to prediction errors between conditions, in regions of interest (ventral striatum, sgACC, vmPFC).

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	t	z
No One PE > Other PE						
Ventral striatum	12	8	-11	2	3.41	3.36
No One PE > Self PE						
No suprathreshold voxels						
Other PE > Self PE + No One PE						
No suprathreshold voxels						
Self PE + No One PE > Other PE						
Ventral striatum	12	8	-11	8	4.52	4.42
sgACC	9	8	-11	2	3.75	3.69
Self PE > Other PE + No One PE						
Ventral striatum	12	11	-11	3	3.34	3.30
No One PE > Self PE + Other PE						
No suprathreshold voxels						

Table S4. *Continued*

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	t	z
Self PE + Other PE > No One PE						
No suprathreshold voxels						
Other PE + No One PE > Self PE						
No suprathreshold voxels						

For all regions, corrected at $p < .05$ FWE-SVC. PE = Prediction error; k =cluster extent. Names of the brain regions were based on the Automated Anatomical Labeling (AAL) atlas.

Table S5. Comparison of whole brain responses to prediction errors between conditions

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	t	z
No One PE > Other PE						
No suprathreshold voxels						
No One PE > Self PE						
No suprathreshold voxels						
Other PE > Self PE + No One PE						
No suprathreshold voxels						
Self PE + No One PE > Other PE						
R Precuneus	15	-100	13	27	4.51	4.23
R Calcarine	18	-100	1		3.7	3.53
L Calcarine	-12	-103	-5	20	3.90	3.70
L Calcarine	-6	-103	1		3.73	3.55
L Postcentral	-36	-34	67	15	3.84	3.65
L Caudate	-12	-7	22	13	3.81	3.63
L Thalamus	-12	-16	13		3.68	3.51
Self PE > Other PE + No One PE						
L Occipital gyrus	-21	-88	22	2385	6.17	5.52
L Cerebellum	-33	-64	-20		5.02	4.64
R Occipital gyrus	24	-88	19		5.00	4.62
L Postcentral	-30	-37	67	701	5.31	4.87
L Precentral	-21	-25	58		4.85	4.50
R Precuneus	6	-43	58		4.63	4.32
R Putamen	33	-13	4	261	4.57	4.27
R putamen	30	-22	4		4.57	4.27

Table S5. *Continued*

Brain Region	*Peak voxel*					
	x	*y*	*z*	*k*	t	z
L Putamen	-30	-13	4	366	4.50	4.21
L Supramarginal gyrus	-57	-28	28		4.15	3.92
R Thalamus	3	-13	55	165	4.49	4.21
L Supplementary motor area	-3	-19	58		4.42	4.14
Median cingulate and paracingulate gyri	6	-1	43		4.13	3.90
R Middle frontal gyrus	36	50	31	22	4.13	3.90
R Supramarginal gyrus	45	-28	34	107	4.30	4.05
R Supramarginal gyrus	54	-28	37		4.02	3.81
R Superior frontal gyrus, dorsolateral	27	-10	70	33	4.25	4.00
L Superior temporal gyrus	-54	-4	4	24	4.04	3.83
L Putamen	-15	11	-11	11	4.02	3.81
R Precental gyrus	42	-10	58	25	3.89	3.69
R Precental gyrus	39	-10	49		3.83	3.64
R Thalamus	15	-22	1	10	3.75	3.57
No One PE > Self PE + Other PE						
No suprathreshold voxels						
Self PE + Other PE > No One PE						
R Precuneus	21	-64	25	99	4.63	4.32
R Calcarine fissure and surrounding cortex	21	-61	16		4.59	4.29
R Cuneus	15	-82	28		3.72	3.54
L Cuneus	-9	-82	22	75	4.33	4.07
L Cuneus	-15	-67	19		4.12	3.90
L Cuneus	-15	-82	31		3.29	3.17
R Insula	36	8	4	58	3.80	3.62
R Insula	36	2	16		3.61	3.46
R Supramarginal gyrus	66	-25	22	61	4.05	3.84
R Superior temporal gyrus	45	-34	22		3.58	3.57
R Supramarginal gyrus	54	-34	31		3.57	3.42
R Heschl gyrus	45	-25	13	18	3.92	3.72
L Middle occipital gyrus	-30	-76	22	12	3.67	3.50
L Middle temporal gyrus	-57	-67	4	10	3.62	3.46
Other PE + No One PE > Self PE						
No suprathreshold voxels						

For all regions, $p < .001$ voxel-level uncorrected, extent-threshold $k = 10$. Names of the brain regions were based on the Automated Anatomical Labeling (AAL) atlas.

Participant instructions

"Welcome! We are going to play a game in the scanner. In this game, you will see two pictures on the screen. You can win or lose points by choosing one of the pictures. If you win, you get +1 point, and if you lose, you get -1 point. But not all pictures are equally good…"

"With both pictures you can win and lose, but with one picture you will win more often, and with the other picture you will lose more often. Try to win as many points as possible! Note: it does not matter whether the picture is on the left or right side of the screen."

"To choose the left picture, you press the left button. To choose the right picture, you press the right button. At the end of the game, you will see how many points you won in total. Your points will be translated to real money using a formula. This amount of money will be paid out to you."

"You will play this game 3 times: for yourself, for another person, and for no one. On the screen, it says for whom you will be playing. Each time you should learn which of the two pictures on the screen is better. Sometimes you play for yourself. When you play for yourself, the gains will be paid out to you."

"Sometimes you play for another person. When you play for another person, the gains will be paid out to another player. This player is someone who participates in this experiment after you. This is a girl or a boy of your age. This person does not know that you are playing for him/her. So, he/she will receive the money you win for him/her without them knowing it is from you. This person will not play the game for you."

"Sometimes you play for no one. When you play for no one, your points don't count and no one will receive your gains."

"Try to respond on time. You will have about 2 seconds to make your choice. We will first do a practice run. Good luck!"

[24 Practice trials (8 per condition)]

"Well done! This was a practice run, so your points don't count yet. In the scanner, we will play the game for real, and you will see at the end of the game how many points you won for yourself and for the other person. Do you have any questions left?"