

This item is the archived peer-reviewed author-version of:

Trust as commodity : social value orientation affects the neural substrates of learning to cooperate

Reference:

Lambert Bruno, Declerck Carolyn, Emonds Griet, Boone Christophe.- Trust as commodity : social value orientation affects the neural substrates of learning to cooperate
Social cognitive and affective neuroscience - ISSN 1749-5016 - (2017), p. 1-9
Full text (Publisher's DOI): <http://dx.doi.org/doi:10.1093/SCAN/NSW170>
To cite this reference: <http://hdl.handle.net/10067/1408050151162165141>

Trust as commodity: social value orientation affects the neural substrates of learning to cooperate.

Running title: **Neural basis of learning to cooperate**

Bruno Lambert^a

Carolyn H. Declerck^a

Griet Emonds^a

Christophe Boone^a

^aACED, faculty of applied economics, university of Antwerp; Prinsstraat 13, 2000 Antwerp, Belgium

Corresponding author: Bruno Lambert, Bruno.Lambert@uantwerp.be, +32 3 265 50 95; ACED, faculty of applied economics, university of Antwerp; S.Z 407, Prinsstraat 13, 2000 Antwerp, Belgium

Abstract

Individuals differ in their motives and strategies to cooperate in social dilemmas. These differences are reflected by an individual's social value orientation: proselfs are strategic and motivated to maximize self-interest, while prosocials are more trusting and value fairness. We hypothesize that, when deciding whether or not to cooperate with a random member of a defined group, proselfs, more than prosocials, adapt their decisions based on past experiences: they 'learn' instrumentally to form a base-line expectation of reciprocity. We conducted an fMRI experiment where participants (19 proselfs, 19 prosocials) played 120 sequential prisoner's dilemmas against randomly selected, anonymous and returning partners who cooperated 60% of the time. Results indicate that cooperation levels increased over time, but that the rate of learning was steeper for proselfs than for prosocials. At the neural level, caudate and precuneus activation were more pronounced for proselfs relative to prosocials, indicating a stronger reliance on instrumental learning and self-referencing to update their trust in the cooperative strategy.

Cooperation, prisoner's dilemma, social preferences, expectation of reciprocity, fMRI

Number of words (excluding abstract, references, tables and figure legends): 5524

Introduction

Greed and fear of betrayal are arguably two of the most important motives that impede cooperation in social dilemmas - situations in which there is a characteristic conflict between self- and collective interest. Ever since Pruitt & Kimmel's seminal paper (Pruitt and Kimmel, 1977) researchers have been trying to identify those factors that account for the willingness to cooperate (to overcome greed) and the expectations of reciprocity (to reduce the fear of betrayal).

The willingness to cooperate has been shown to be highly heterogeneous among individuals and reflects a person's social value orientation (SVO; Van Lange, 2000). Individuals with a prosocial value orientation have other-regarding preferences, prefer equal outcomes, and cooperate readily in social dilemmas because they have internalized a moralistic, cooperative norm (Bogaert *et al.*, 2008). In contrast, individuals with a proself value orientation have self-regarding preferences, maximize self-interest in social dilemmas and will therefore defect by default. However, there are many reports that proselfs will cooperate when there are extrinsic incentives that align self-interest with collective interest. This accomplishes a goal transformation whereby greed is no longer an obstacle to cooperation, leading proselfs to strategically cooperate (De Cremer and Van Vugt, 1999). This is, for example, the case when dyadic interactions are repeated, allowing profits for both parties involved to accumulate over time (Axelrod and Hamilton, 1984), when cooperation yields synergy by changing the pay-off structure from a mixed motive dilemma to a coordination task (Boone *et al.*, 2010), or when one's reputation is at stake (Simpson and Willer, 2008; Declerck *et al.*, 2014).

fMRI studies furthermore corroborate that proselfs are more strategic and calculative when they make decisions, while prosocials are more willing to cooperate because they conform to social norms. When participants under the fMRI scanner were playing a series of social dilemma games with different incentive structures (Emonds *et al.*, 2011), only proselfs adapted their behavior in accordance with incentives, and this was accompanied by increased activation in the dorsolateral prefrontal cortex, a region implicated in cognitive control and cost/benefit analysis (Miller and

Cohen, 2001). Proselfs also showed more activation in the precuneus which is an important region in self-centered cognition (Kircher *et al.*, 2000; den Ouden *et al.*, 2005; Cavanna and Trimble, 2006) and the posterior superior temporal sulcus (pSTS), a region involved in comparing intentions of self-versus others (Saxe and Wexler, 2005) and hence a crucial element in maximizing pay-offs for self. In contrast, prosocials showed more activation in the anterior portion of the STS, associated with routine, moral judgments (Borg *et al.*, 2006) which suggests that they are more norm-compliant. In addition, when prosocials are treated unfairly, they show more amygdala and nucleus accumbens activity, even when they are under cognitive load (Haruno *et al.*, 2014). This further suggests that prosocials engage more in automatic decision making.

The influence of SVO has received less attention with respect to the second factor that determines cooperative decision-making, namely the expectations of reciprocity. According to the triangle hypothesis (reviewed in Bogaert *et al.* (2008)), proselfs are more likely to assume others are also proselfs, whereas prosocials have a more heterogeneous view of the social world. Accordingly proselfs would generally expect little reciprocity from game partners, while prosocials would rely on trust in order to decide whether or not to cooperate. This is corroborated by an experimental study showing that prosocials are indeed more likely to cooperate in a one-shot social dilemma game when they either have high dispositional trust, or when they have a chance to familiarize themselves with the game partner (Boone *et al.*, 2010). For proselfs, dispositional trust does not matter with respect to the decision to cooperate, and they are more likely to abuse trust signals if this is to their advantage (Emonds *et al.*, 2014).

However, there are many situations in which it does pay-off to trust partners. Whenever we find ourselves in new and transient groups that impose long-term collaboration, cooperation may be the best strategy leading to the most lucrative outcome, provided that there are sufficient other individuals in the group that are also cooperative. This is the case when we start to work in a new firm, join a sports club, or even purchase on e-bay. In these settings trust is a commodity that makes

it possible to maximize self-interest. Such “instrumental trust” could be acquired through reinforcement learning processes by which the base-rate of reciprocity in the group of interacting parties is established. We propose that “learning to trust” in such situations is a strategy by which proselves will adapt their rate of cooperation to the expected level of reciprocity.

To test this hypothesis, we conduct an fMRI study assessing the neural substrates of decision-making in a sequential prisoner’s dilemma (PD) game where a cooperative decision of the first mover is only determined by the expectation of reciprocity and not by greed. We investigate if prosocials and proselves differ in the underlying mechanism by which they are forming expectations of reciprocity in a transient group of anonymous partners and adapt their level of cooperation accordingly. If proselves, more than prosocials, update their level of trust based on reinforcement learning, we expect this to be reflected by a relative greater increase in activation of the caudate nucleus, a subcortical region of the brain implicated in instrumental learning (O’Doherty *et al.*, 2004) and updating behavior (King-Casas *et al.*, 2005; Baumgartner *et al.*, 2008; Waegeman *et al.*, 2014).

The sequential PD lends itself well to studying how expectations of reciprocity are formed. For the first mover there is an incentive to cooperate because he or she can potentially earn more by cooperating, but this is contingent on the second player’s decision to reciprocate. Because the decision of the first player is revealed to the second player, the second player cannot fare worse than the first player because a defect decision of the first player is unlikely to be positively reciprocated. This removes greed as a motive for the first player. By repeating the sequential PD interactions within a closed group of anonymous partners that return randomly, we simulate the occurrence of real-life transient groups in which cooperation can emerge because possible future interactions “cast a shadow back on the present and thereby affect the current strategic cooperation” (Axelrod and Hamilton, 1984). In this setting it is possible for the first mover to establish the base-rate of reciprocity by relying on instrumental learning processes whereby each instance of positive reciprocation reinforces the presumption that cooperation is paying-off.

From agent-based simulations (Riolo *et al.*, 2001) and laboratory experiments (Efferson *et al.*, 2008) we know that cooperation can emerge in large pools of anonymous partners that are randomly matched, and typically a stable equilibrium emerges with roughly 60% cooperators (Fischbacher *et al.*, 2001; Bowles and Gintis, 2004; Kurzban and Houser, 2005; Balliet *et al.*, 2009). Similarly, the experiment of Kiyonari *et al.* (2000) indicated that second movers in a sequential PD reciprocated cooperation 62% of the times, which was much higher than the 38% cooperation in the simultaneously played PD. A plausible interpretation for this jump is that the “reciprocal exchange” nature of the game, which is more salient in the sequential than in the simultaneous PD, provides a strategic incentive to cooperate (Kiyonari *et al.*, 2000), and this would be especially valuable for prosocial individuals (Simpson, 2004).

In summary the main hypothesis we test in this current study is that, given 60% reciprocation in a repeated and sequentially played PD, “learning to cooperate” will be more pronounced for prosocials, and this will be reflected in the neural mechanisms (especially the caudate nucleus) that substantiate instrumental learning.

Methods

Participants

We recruited participants via e-mail, flyers, and web-based advertisements in which the study was introduced as an investigation of the brain areas that become activated during economic decision-making. We selected thirty-eight participants (22 females, average age = 24.6, S.D. = 4.5) based on (i) a medical screening questionnaire, to make sure the participant met all the safety criteria for MRI examination, and (ii) their SVO, assessed using the decomposed method (Van Lange *et al.*, 1997). This measure consists of nine items in which the participant can choose between three distributions of points allocated to oneself and an anonymous partner. Each of these distributions represents a particular social value orientation. We selected only those participants with a consistent prosocial or

individualistic orientation (i.e., at least six out of the nine choices in the SVO questionnaire were consistent with that orientation). We refer to the individualistic orientation as “proself.”

All procedures were approved by the medical ethics commission of the University of Antwerp. Debriefing occurred at the conclusion of the study by contacting participants by email and referring them to a website where the intent, results, and procedures of the experiment were fully explained.

Paradigm

Participants played the role of the first mover in a repeated and sequentially played PD with a number of different but returning partners while under the MRI scanner. A pay-off matrix of a PD game played between two people is shown in Figure 1. When the game is played sequentially, the first mover can potentially earn more by cooperating, but this is contingent on the second mover’s decision to reciprocate. If the second mover defects, the first mover loses more than if he/she had not cooperated. Thus the participant’s intent to cooperate in this game should be a function of the expected reciprocity rate of the partner pool. In this experiment, there were supposedly 25 partners and their decisions were, unbeknownst to the participant, computer-programmed to reciprocate cooperation in 60% of the cases. The program was such that there was never a cooperative decision after the participant chose to defect.

		Me	
		C	D
Person 18	C	5 / 5	0 / 6
	D	6 / 0	1 / 1

Figure 1. A PD pay-off matrix as shown in the experiment. The participant can choose between the C or D column. If the participant chooses C (the cooperative decision) 5 points will be earned if the

partner subsequently reciprocates by also choosing C (mutual cooperation); 0 points will be earned if the partner were to choose D. If the participant chooses D, the partner response is set to always be D, in which case the participant earns 1 point. In the actual experiment, we changed 'C' and 'D' into 'K' and 'L' to avoid bias.

To increase the participants' believability that they were playing against real partners, they were first shown 25 pictures representing the alleged, anonymous partners. These pictures were obtained from people who previously participated in similar experiments and who approved that their pictures would be used for this purpose. Participants were told that, for each of the 120 trials they were to play under the scanner, they would be randomly matched with one of these 25 partners. Furthermore, they were made clear that their profit earned in each trial would depend upon the combination of their own choice and the choice made by the partner. During the game, the presumed partner was always identified with a number and not with the previously seen picture. Because these numbers corresponded to 25 possible partners, it was nearly impossible to keep track of returning partners.

Each participant received written instructions explaining that they were to play the first mover in a series of 120 PD games. To make sure all participants understood the pay-offs of the PD correctly, they had to answer three test questions correctly before starting the actual experiment. Once in the scanner, the 120 PD games were played in six rounds of twenty trials each, with a short break after each round. Within a round, participants could not be matched more than once with a certain partner.

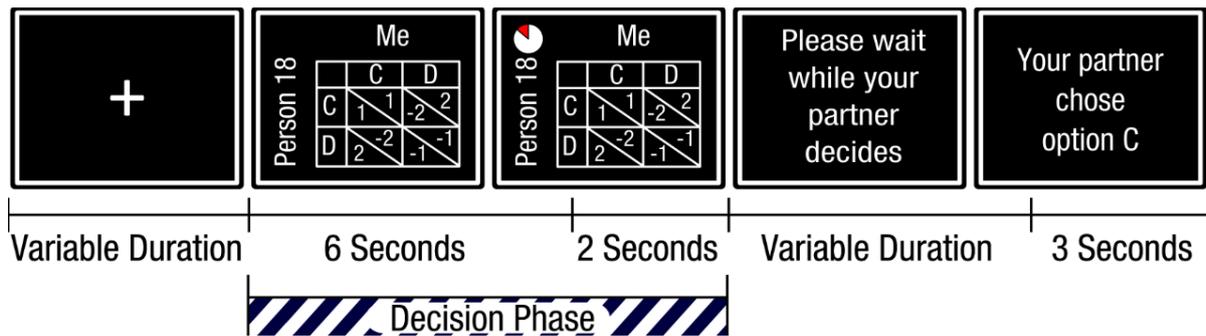


Figure 2. Time line of the experiment. A fixation cross marks the inter-trial interval which lasts between 1.5 and 4 seconds, after which a screen depicting the PD appears. Six seconds later, an indicator appears and stays on the screen for an additional 2 seconds, signaling the end of the participants' allotted time to make a decision. The decision phase can end earlier when one of the buttons is pressed. In between the decision phase and the feedback phase a screen appears for 2 to 4 seconds asking the participant to wait. The feedback phase starts when the screen revealing the partner response is shown, which lasts 3 seconds.

The time sequence of the experiment is shown in Figure 2. To avoid boredom, 25 different PD matrices were used, and no matrix is used more than once in a given round. The difference in pay-off between defection or cooperation after a cooperative decision was kept equal for each PD. The values in the game matrix represent points to be exchanged for money at the end of the experiment (1 point = € 0.02; average earnings: € 9.32 + € 10.00 show-up fee). Participants indicated their choice (to cooperate or not) by pressing the corresponding button on the response box held in their right hand (Lumia model LU400, Cedrus, CA, USA). The allotted 8 seconds per trial proved sufficiently long, as the responses were successfully registered in 99% of the cases, and no-one missed more than six out of the 120 decisions. Stimulus presentation and response logging was conducted with the Presentation® software (Neurobehavioral Systems, Inc, Albany, CA, USA).

This experiment was part of a larger study comprising two independent and different experiments using the same participant pool (the results of the first experiment are published elsewhere, see

Emonds *et al.* (2014)). Because the participants received no feedback after the first experiment, it is highly unlikely that this would have affected the second experiment.

Modeling the learning effect

To test if participants are learning to trust and update their cooperation rate according to the 60% reciprocity feedback, we adopted two different methods. First we examined how behavior changes *over time* and plot the number of cooperative decisions as a function of trial number. Second, we tested how individuals update their decisions in function of their accumulated experiences in the game, using the *Experience weighted attraction (EWA) learning model* developed by Camerer and Hua Ho (1999). This model (explained in more detail in the supplementary material) computes attractions for each strategy that can be used in a game context (i.e., cooperate or defect in the prisoner's dilemma), and this attraction is updated after each trial (see eq. 1 in the supplementary material). The attraction for cooperation is calculated by taking the sum of (i) a first term that accounts for previous attractions towards cooperation during the game, weighted by a free discounting parameter that weighs recent trials more heavily than earlier ones, and (ii) a second term that takes the actual or forgone payment of the trial in consideration. This sum is then normalized by dividing it by the discounted number of encounters that have taken place. To investigate how individuals adapt their behavior in function of how much they learned from their previous experience, it is the first term of the EWA equation that is of interest.

fMRI image acquisition

Images were collected with a 3 Tesla Siemens Trio scanner and an 8-channel head coil (Siemens, Erlangen Germany). A T1-weighted magnetization-prepared rapid acquisition with gradient echo (MP-FRAGE) protocol was used to create anatomical images (256 × 256 matrix, 176 0.9 mm sagittal slices, field of view (FOV) = 220 mm). Functional images were acquired using T2*-weighted echo planar imaging (EPI) (repetition time (TR) = 2400 ms, echo time (TE) = 30 ms, 64 × 64 image

resolution, FOV = 224 mm, 39 3 mm slices, voxel size = 3.5 × 3.5 × 3.0 mm). Due to technical problems during the scanning procedure, the results of one participant was lost, leaving a total of 19 proselves and 18 prosocials for image analysis.

fMRI data analysis

Image analysis of the 37 brain volumes (39 slices with a TR of 2.4s) was conducted with Matlab (MATLAB and Statistics Toolbox Release 2012a, The MathWorks, Inc., Natick, MA, USA) and the Statistical Parametric Mapping package (SPM8; Wellcome Department of Cognitive Neurology, London, UK).

Three general linear models (GLMs) were created for each participant. In each case, the blood oxygen level dependent (BOLD) signal was the dependent variable. The event of interest was the decision phase, defined as the time interval between the appearance of the slide depicting the PD matrix and the participant's response (Figure 2). We did not differentiate between cooperation or defect decisions because we are interested in revealing the underlying neural correlates of *learning* which depends on previous experiences with both strategies. The actual decision is not important in this setting: one can learn to cooperate *or* defect, depending on the type of feedback that is given.

With the first GLM, we explore which brain regions are involved in the initial decision making process of the first mover in the sequentially played prisoner's dilemma game, and we test if they differ between prosocials and proselves who we presume have different underlying motives to cooperate. Here we consider only the decisions made during the first round (i.e. first 20 trials), because this is where we expect the difference in the intrinsic motivation to cooperate to differ between prosocials and proselves. It is also the round in which in the difference in level of cooperation between prosocials and proselves is the most significant (see results section).

With the second and the third GLM we investigate the neural correlates of learning (and how they differ between prosocials and proselfs), relying on the entire course of the experiment (i.e., the full 120 trials). We model “learning to cooperate” in two different ways: by adding the trial number as parametric modulator (GLM 2), and by adding the learning term of the EWA model (GLM 3; see eq. 3 in the supplementary material).

For all three GLMs, all regressors were convolved with the haemodynamic response function. Six movement parameters were added to account for head movement in six dimensions. For each of the three models, we started by conducting a whole brain first level analysis for the contrast [decision – baseline] before conducting a second level analysis which further investigates possible interaction effects with SVO. For subsequent region of interest (ROI) analyses, we relied on independent coordinates derived from the maximum probability atlas of the human brain (Hammers *et al.*, 2003) and from the AAL atlas (Tzourio-Mazoyer *et al.*, 2002).

Results

Behavioral data

For each group (proselfs and prosocials), we plotted the proportion of cooperative decisions in each of the 120 trials (Figure 3). This reveals that proselfs start out cooperating at a lower level than prosocials (proportion of cooperation in first 20 rounds: mean proselfs = 0.50, 95% CI [0.45 0.55]; mean prosocials = 0.62, 95% CI [0.57 0.67], p -value < 0.001). Their behaviors converge during the experiment (proportion of cooperation in last 20 rounds: mean proselfs = 0.70, 95% CI [0.66 0.75]; mean prosocials = 0.71, 95% CI [0.66 0.76]). Furthermore, the overall proportion of cooperation increases throughout the experiment, and this learning effect is more pronounced for proselfs. To verify this statistically, we estimated a logistic regression model with random effects to account of the unobserved differences among the 38 participants with cooperative decision (coded 1, defect coded 0) in each of the 120 trials as the dependent variable. This corroborates that there is a

significant effect of trial number ($B = 0.010$, 95% CI [0.008 0.013], p -value < 0.000 , Wald $\chi^2 = 79.50$) but not of SVO (coded 0 for prosocials and 1 for proselfs). When adding the interaction term SVO * trial number to the regression model, we observe that the increase in cooperative decisions differs significantly for prosocials and proselfs ($B = 0.0066$, 95% CI [0.002 0.011], p -value = 0.005, Wald $\chi^2 = 86.07$), indicating that indeed proselfs are adapting their behavior over time more so than prosocials. We repeated this analysis with the EWA estimated probability of making a cooperative decisions instead of the observed decisions, and obtain very similar results (see supplementary material and Figure S1).

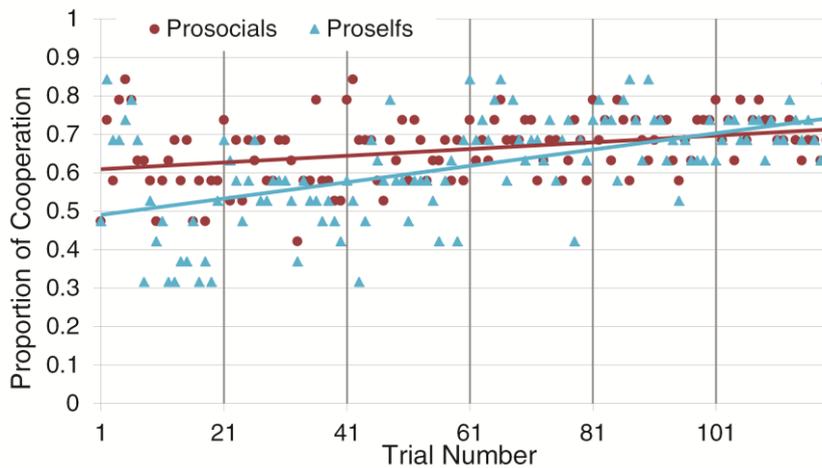


Figure 3. Proportion cooperative decisions made by the participants for each of the 120 trials, plotted for prosocials and proselfs. Best fit lines, based on least squares, are shown.

If prosocials have internalized a cooperative norm (making them less reliant on instrumental learning), we would expect them to respond more intuitively than proselfs, which would be noticed in shorter response latencies. Therefore we plotted the average reaction times (RT) for proselfs and prosocials (Figure 4). Consistently, we observed that the reaction time is faster for prosocials than proselfs (Mean prosocials = 2.60 s, 95% CI [2.57 s 3.79 s]; Mean proselfs = 3.18 s, 95% CI [2.09 s 3.11 s]). A linear regression with random effects with RT as the dependent variable shows a marginally significant effect of SVO ($B = -0.58$, 95% CI [-1.27 0.11], p -value = 0.1), and a significant effect of trial number ($B = -0.01$, 95% CI [-0.011 -0.0091], p -value < 0.001 , Wald $\chi^2 = 413.20$). The interaction term

SVO * trial number is also significant ($B = 0.0031$, 95% CI [0.001 0.005], p -value = 0.002, Wald $\chi^2 = 424.13$), revealing that prosocials respond faster, and proselfs slower, especially early on in the experiment. Thus as time progresses, proselfs show less need to deliberate their choice. This interaction between SVO and time also indicates that the learning effect is greater for proselfs.

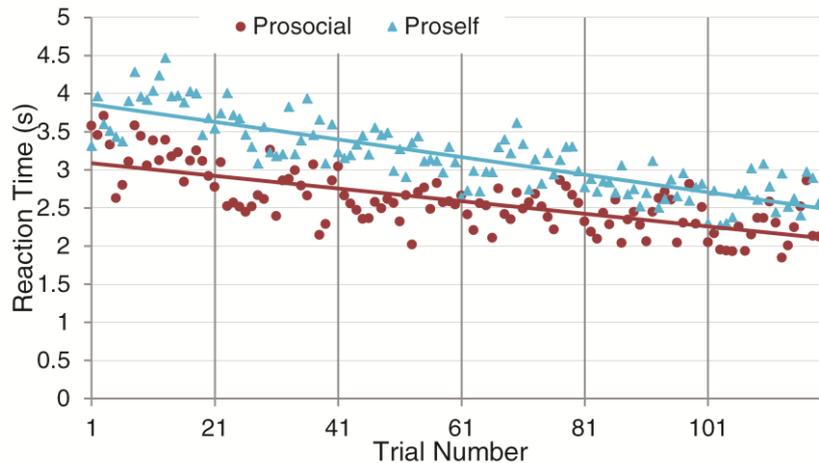


Figure 4. Average reaction times for prosocials and proselfs throughout the experiment. Lines of best fit are based on least squares.

Functional MRI data

We first contrasted the decision phase with baseline activation during the first round (average of 20 trials) in the whole brain (using the first GLM). Clusters of significant activation (family wise error [FWE] corrected p -value < 0.05, whole brain) are reported in Table 1. We note that there is significant activation in the precuneus and insula, two regions of great interest in social decision making. Activation of the insula has been reported as a necessary component of overcoming betrayal aversion (Aimone *et al.*, 2014), while the precuneus is involved in self-referencing (Cavanna and Trimble, 2006). The latter region would be especially important to proselfs (as shown in Emonds *et al.* (2011)), who we hypothesize are trusting instrumentally, because mutual cooperation is in their best interest. However, a subsequent whole brain analysis contrasting proselfs and prosocials during the decision phase in the first round did not yield any significant differences in brain activation,

neither did an ROI on the precuneus (coordinates derived from Tzourio-Mazoyer *et al.* (2002)) and the insula (coordinates derived from Hammers *et al.* (2003)).

Table 1.

Clusters of voxels significantly activated during the decision phase in the first round (first 20 trials)

Region^a	BA	Side	x	y	z	Size^b	t	p
Frontal								
Middle Frontal Gyrus	-	L	-36	-2	54	133	7.65	<0.001
	6	R	34	2	64	121	8.24	<0.001
Inferior Frontal Gyrus	9	L	-42	2	30	271	8.13	<0.001
Parietal								
Precuneus	-	L	-8	-66	58	3062	10.50	<0.001
Occipital								
Middle occipital Gyrus	-	L	-36	-64	-8	429	8.06	<0.001
Lingual Gyrus	-	R	10	-82	-12	373	8.80	<0.001
Limbic								
Cingulate Gyrus	-	R	6	18	44	372	10.52	<0.001
Sub-lobar								
Insula	-	L	-38	16	2	193	7.84	<0.001
	-	R	34	24	2	346	12.69	<0.001

Summary of all significantly (whole brain FWE corrected p-value < 0.05) activated clusters during the decision phase (coordinates are in MNI space). Only clusters with more than 100 voxels are shown.

BA = Brodmann area, L = left, R = right.

^aRegions were determined using the AAL atlas (Tzourio-Mazoyer *et al.*, 2002).

^bNumber of statistically significant voxels (voxel size of 2.0 x 2.0 x 2.0 mm).

To investigate learning, the second and third GLM took the decisions of each of the 120 trials into account. In the second GLM, we added the trial number as a parametric modulator and contrasted this with baseline. Results (FWE corrected p-value < 0.05, whole brain) are listed in Table 2. As hypothesized, we observe significant activation in the caudate nucleus, which has previously been identified as one of the most important regions implicated in instrumental learning and updating behavior (O'Doherty *et al.*, 2004; King-Casas *et al.*, 2005; Baumgartner *et al.*, 2008; Waegeman *et al.*, 2014).

Table 2.

Clusters of voxels significantly modulated by trial number during the decision phase in the entire experiment (120 trials)

Region^a	BA	Side	x	y	Z	Size^b	t	p
Frontal Lobe								
Precentral Gyrus	-	L	-14	-32	72	553	8.43	<0.001
Parietal Lobe								
Precuneus	-	L	-10	-48	62	188	7.39	<0.001
Sub-Gyral	-	R	22	-58	56	927	8.30	<0.001
Limbic Lobe								
Parahippocampal Gyrus	-	L	-24	-44	2	101	7.56	<0.001
Sub-lobar								
Caudate	-	L	-14	20	4	2157	9.65	<0.001
Insula	-	R	32	-22	14	107	8.53	<0.001

Note: Summary of all clusters where activation during the decision phase was significantly (whole brain FWE corrected p-value < 0.05) modulated by trial number (coordinates are in MNI space). Only clusters with more than 100 voxels are shown. BA = Brodmann area, L = left, R = right.

^aRegions were determined using the AAL atlas (Tzourio-Mazoyer *et al.*, 2002).

^bNumber of statistically significant voxels (voxel size is 2.0 x 2.0 x 2.0 mm).

When contrasting proselfs versus prosocials in this whole brain contrast, no activation survived correction (FWE corrected p -value < 0.05). We therefore perform ROI-analyses on the caudate (derived from Hammers *et al.* (2003)), the precuneus, and the insula. Figure 5 shows the β values for the parametric modulation of trial number in these regions (GLM 2), revealing a significant greater effect for proselfs (compared to prosocials) in the former two. Figure 6 projects these significantly activated clusters on a standard brain for both the proself (Figure 6a) and prosocial (Figure 6b) group.

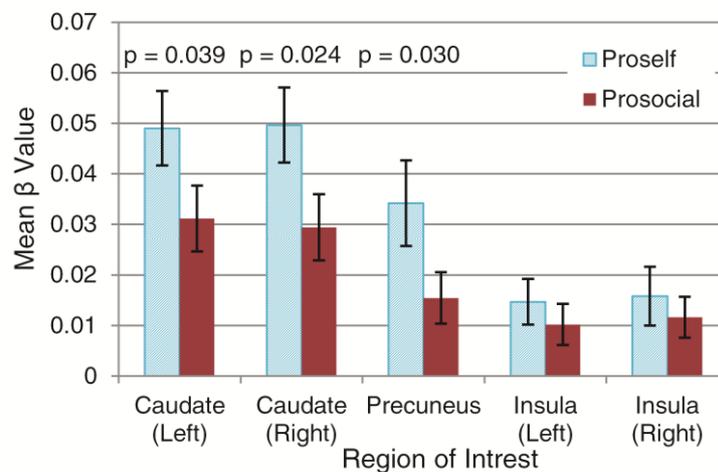


Figure 5. Parameter estimates for brain activation modulated by trial number during the decision phase in three ROI's. Error bars represent Standard Errors of the Mean. P-values of t-tests reveal significant differences in activation between the prosocial and proself group in the caudate and the precuneus, but not in the insula.

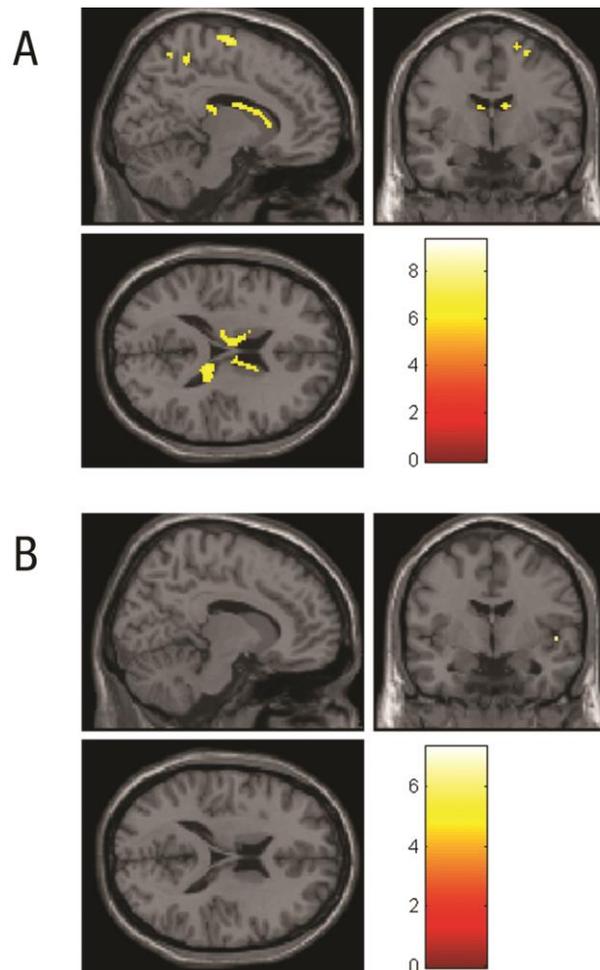


Figure 6. Whole brain analysis showing activation (in sagittal, coronal and transvers section; $x = 12$, $y = -4$, $z = 18$) modulated by the trial number during the decision phase for (A) 19 proselves and (B) 18 prosocials. T value cutoff = 5.81 (corresponding to an FWE corrected p-value = 0.05).

We repeated the same whole brain and ROI analyses using the third GLM, which includes the learning term derived from the EWA model as a parametric modulator. The whole brain analysis does not reveal any regions that are significantly affected by experience learning. The ROI analyses, however, show that the effect of this parametric modulator on the BOLD signal in the caudate and precuneus differ significantly between prosocials and proselves, with prosocials showing relatively less activation in these regions than proselves (see Figure 7). This is consistent with the hypothesis that adopting a cooperative strategy based on previous experience (learning to trust) is less instrumental (i.e. less dependent on reinforcement and self-referencing) for prosocials compared to proselves. For

prosocials, caudate and precuneus neural activity is negatively associated with experience based attraction towards cooperation. Thus, when deciding to cooperate, they seem to rely less on those regions, given their imprinted preference to seek out mutual cooperation. Updating their attraction to cooperation, conditional on the partner’s strategy, is therefore likely to occur by a different mechanism that does not rely as much on either reinforcement learning (involving the caudate nucleus) or self-referencing (involving the precuneus).

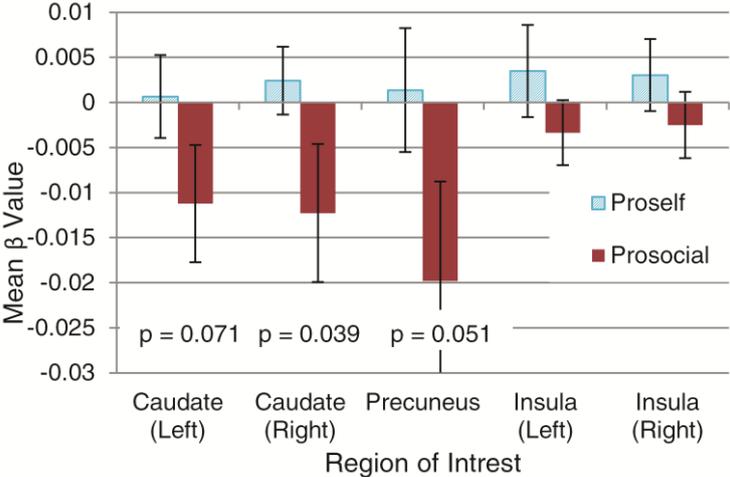


Figure 7. Parameter estimates for brain activation modulated by the first term of the EWA model during the decision phase in three ROI’s. Error bars represent Standard Errors of the Mean. P-values of t-tests reveal significant differences in activation between the prosocial and proself group in the caudate and the precuneus, but not in the insula.

Discussion

The results contribute to the body of knowledge indicating that individuals with different social value orientations rely on different strategies to cooperate in social dilemmas. In line with previous findings the behavioral data show that, in a repeated sequential PD, first mover prosocials have an overall greater willingness to cooperate (they start out cooperating at a higher rate and they have faster reaction times). Proselfs, in this setting, are learning to cooperate because it is a strategy that

pays off in the long run; their behavior converges with that of prosocials towards the end of the experiment.

The fMRI data shed light on the underlying neural mechanism that may drive some of the difference in behavior between prosocials and proselfs. The initial greater willingness to cooperate of prosocials (relative to proselfs) was, in the current study, not associated with any notable difference in the pattern of brain activation. When it comes to reinforcement-based learning, however, the data are consistent with the hypothesis that proselfs (more than prosocials) are learning to cooperate by adapting their expectations of reciprocity instrumentally, relying on the caudate and precuneus.

As hypothesized, proselfs showed a substantially greater increase in neural activation in the caudate nucleus over time (Figure 5). Furthermore, for proselfs, but not for prosocials, the caudate nucleus remains active when the attraction to the cooperative strategy is updated based on the accumulated (positive and negative) feedback throughout the experiment. The caudate nucleus has previously been shown to be the key structure involved in establishing stimulus-response contingencies (Packard and Knowlton, 2002; O'Doherty *et al.*, 2004). It becomes more active with perseveration and therefore plays an important role in the decision to either update behavior or maintain the status quo (Baumgartner *et al.*, 2008; Smith-Collins *et al.*, 2013; Waegeman *et al.*, 2014). In a repeated dyadic trust game, the caudate nucleus is reported to become increasingly activated as partners are gaining more experience with each other, with peak activity reported at the moment the investor anticipates positive reciprocation (King-Casas *et al.*, 2005). This instrumental trust by which expectations of reciprocity are formed is, in the current study, more pronounced for proselfs. For prosocials in the current experiment, caudate nucleus activation actually decreased with increasing experience (Figure 7), suggesting that trust is not learned instrumentally, but perhaps an intrinsic feature that is adopted automatically from the start.

In addition to the differential activation of the caudate nucleus, the data also reveal that, compared to prosocials, proselfs show a greater increase in precuneus activation as time progresses (Figure 5)

and a relative greater precuneus involvement in learning (Figure 7). The precuneus is activated when comparing the outcome of decisions for self and others, and plays an important role in solving social dilemma problems, especially for proselves who rely more heavily on self-referencing to compute the strategy with the highest pay-off (Emonds *et al.*, 2011). Again, it makes sense that this region would become more active with time, and that activation of this region remains active for proselves as they gain experience with the cooperative strategy: only at the end of the experiment do the proselves have sufficient ground to establish positive expectations of reciprocity. At the onset of the experiment, they have no basis for comparison yet, and the role of the precuneus may be less salient (Emonds *et al.*, 2014).

In addition to the caudate and precuneus, the data reveal that the insula also becomes increasingly activated as time progresses, and this is equally true for prosocials as well as proselves (Table 2). The insula has a well-established role in emotion-processing, and its activation in the context of solving dilemma-type problems points to automatic processing (Kuo *et al.*, 2009). It is also activated when overcoming betrayal aversion (Aimone *et al.*, 2014), which is consistent with the current finding that decision latencies are decreasing with time and, that betrayal aversion is also steadily decreasing as positive expectations of reciprocity are formed. Considering that betrayal aversion is an important reason why prosocials might chose not to trust, it is not surprising that the increase in insula activation does not differ between prosocials and proselves (see Figures 5 and 7).

These data, revealing caudate, precuneus, and insula activation in learning to cooperate, are particularly relevant in the light of a recent experiment by Watanabe *et al.* (2014). These authors scanned participants that were supposedly embedded in a chain of reciprocal donations. The participant had to decide whether or not to donate a sum of money to the next individual in the chain (which was costly for the participant but beneficial for the other). In the first condition, the participant had received money from the previous person in the chain and decided to “pay it forward” to the next one. This activated the caudate nucleus together with the insula, pointing to the

involvement of emotions in cooperative decision making. In the second condition, the participant (regardless of what he or she had received previously) decided to donate money to the next person in the chain knowing that this person had a history of donating. In this “reputation-based” condition, the caudate together with the precuneus were activated. Interestingly, we find that these two regions (precuneus and caudate) are more activated in proselves, which is consistent with their strategic nature: if they are truly establishing the base-rate of reciprocity to solve repeated PD games, it is important that they learn the history of the cooperative behavior of returning partners.

Finally, the current findings also extend the results described by Smith-Collins *et al.* (2013). These authors investigated how participants make decisions in a trust game where the same and new partners re-appear in consecutive rounds and feedback is given. They report that activation in the caudate is associated with successfully adapting behavior after being confronted with unexpected cooperation or betrayal, i.e. trusting after unexpected cooperation or distrusting after unexpected betrayal. While Smith-Collins *et al.* (2013) focus on the response towards an *identifiable* partner when group composition is changing, we show that trust-based learning (with increasing caudate activity over time) can also take place based on expectations of the entire group.

We also note limitations to the study. First fMRI data can be influenced by elements that are unrelated to the decision-making process, such as scanner drift (which tends to increase neural activity over time), fatigue, or other effects that are only marginally related to the task. These extraneous factors call for caution in the interpretation of the results, but they cannot explain the differences that we observe between two groups that received the exact same treatment and only differed in personality.

A second limitation pertains to the experimental design and challenges the generalizability of the conclusions. While we decided *a priori* on 60% positive reciprocity (the level needed to sustain mutual cooperation in natural populations, see introduction), this rate also constraints what can be learned. Prosocials are more likely to cooperate when they sense their partners are cooperative

(Kuhlman and Marshello, 1975), so with a 60% cooperation rate, there is little left to be learned for them because their expectation of reciprocity is easily met (i.e., the feedback they receive in early rounds is sufficient to allow their cooperative norm to surface). Hence the current experimental design cannot disentangle whether prosocials are not activating brain regions involved in instrumental learning and self-referencing because they truly differ in the way they integrate newly accumulated information in their utility function, or because their preconceived expectations of reciprocity did not differ substantially from reality. In contrast, proselfs, who are not inclined to cooperate naturally, still have a lot to learn before they can establish that trust is a self-serving strategy, and the data show the involvement of the caudate and precuneus. Hence at least for proselfs, learning to trust is a commodity that they have acquired instrumentally.

An interesting avenue for future studies is to investigate if, and how, prosocials and proselfs differ when they are “unlearning to cooperate,” if reciprocity rates were 40% or lower. As mutual cooperation is potentially the most lucrative outcome for the first player in a sequential prisoner’s dilemma game, we expect that proselfs would still need to rely on the same instrumental learning and self-reference processes to establish that the reciprocity rate is too poor to adopt a cooperative strategy. The interesting question is how prosocials would adapt their behavior in a pool of negative reciprocators. Being behavior assimilators, they are likely to also adapt to a defect strategy, but we do not know whether this change in behavior can be attributed to the same instrumental learning mechanism as proselfs. Given prosocials’ strong egalitarian orientation and retaliatory nature, “unlearning to cooperate” might reveal itself more in a neural signature that updates decision making emotionally rather than instrumentally. This would be consistent with findings that the breaches of fairness that cause punitive actions by prosocials are driven by amygdala and nucleus accumbens activation (Haruno *et al.*, 2014).

In sum, this is the first study to report on the neural basis by which individuals with different social values are forming expectations of reciprocity and thereby learning to cooperate when they

experience 60% reciprocation in a transient pool of anonymous partners. While prosocials' decisions based on expectations are more automatic and change little over time, proselfs are learning to trust instrumentally, whereby they activate (more than prosocials) the caudate nucleus and the precuneus.

Funding

This research was funded by an ID-BOF grant (5340) from the University of Antwerp.

Author Contributions

G. Emonds, C. Declerck and C. Boone developed the study conceptually and methodologically. Data collection was performed by G. Emonds. Data analysis was conducted by B. Lambert. The paper was written with the concerted effort of B. Lambert, C. Declerck, and C. Boone. All authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the - publication of this article.

References

- Aimone, J. A., Houser, D. & Weber, B. (2014) Neural signatures of betrayal aversion: an fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences*, 281(1782).
- Axelrod, R. & Hamilton, W. D. (1984) *The evolution of cooperation*, New York, Basic Books.
- Balliet, D., Parks, C. & Joireman, J. (2009) Social value orientation and cooperation in social dilemmas: a meta-analysis. *Group Processes & Intergroup Relations*, 12(4), 533-547.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U. & Fehr, E. (2008) Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*, 58(4), 639-650.
- Bogaert, S., Boone, C. & Declerck, C. (2008) Social value orientation and cooperation in social dilemmas: a review and conceptual model. *British Journal of Social Psychology*, 47, 453-480.
- Boone, C., Declerck, C. H. & Kiyonari, T. (2010) Inducing cooperative behavior among proselves versus prosocials: the moderating role of incentives and trust. *Journal of Conflict Resolution*, 54(5), 799-824.
- Borg, J. S., Hynes, C., Van Horn, J., Grafton, S. & Sinnott-Armstrong, W. (2006) Consequences, action, and intention as factors in moral judgments: an fMRI investigation. *Journal of Cognitive Neuroscience*(18), 803-817.
- Bowles, S. & Gintis, H. (2004) The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology*, 65(1), 17-28.
- Camerer, C. & Hua Ho, T. (1999) Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4), 827-874.
- Cavanna, A. E. & Trimble, M. R. (2006) The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129, 564-583.
- De Cremer, D. & Van Vugt, M. (1999) Social identification effects in social dilemmas: a transformation of motives. *European Journal of Social Psychology*, 29(7), 871-893.
- Declerck, C. H., Boone, C. & Kiyonari, T. (2014) No place to hide: when shame causes proselves to cooperate. *Journal of Social Psychology*, 154(1), 74-88.

- den Ouden, H. E. M., Frith, U., Frith, C. & Blakemore, S. J. (2005) Thinking about intentions. *Neuroimage*, 28(4), 787-796.
- Efferson, C., Lalive, R. & Fehr, E. (2008) The coevolution of cultural groups and ingroup favoritism. *Science*, 321(5897), 1844-1849.
- Emonds, G., Declerck, C. H., Boone, C., Vandervliet, E. & Parizel, P. M. (2011) Comparing the neural basis of decision making in social dilemmas of people with different social value orientations, a fMRI study. *Journal of neuroscience, psychology, and economics*, 4(1), 11-24.
- Emonds, G., Declerck, C. H., Boone, C., Seurinck, R. & Achten, R. (2014) Establishing cooperation in a mixed-motive social dilemma. An fMRI study investigating the role of social value orientation and dispositional trust. *Social Neuroscience*, 9(1), 10-22.
- Fischbacher, U., Gächter, S. & Fehr, E. (2001) Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3), 397-404.
- Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., et al. (2003) Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping*, 19(4), 224-247.
- Haruno, M., Kimura, M. & Frith, C. D. (2014) Activity in the nucleus accumbens and amygdala underlies individual differences in prosocial and individualistic economic choices. *Journal of Cognitive Neuroscience*, 26(8), 1861-1870.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R. & Montague, P. R. (2005) Getting to know you: reputation and trust in a two-person economic exchange. *Science*, 308(5718), 78-83.
- Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., et al. (2000) Towards a functional neuroanatomy of self processing: effects of faces and words. *Cognitive Brain Research*, 10(Article), 133-144.
- Kiyonari, T., Tanida, S. & Yamagishi, T. (2000) Social exchange and reciprocity: confusion or a heuristic? *Evolution and Human Behavior*, 21(6), 411-427.

- Kuhlman, D. M. & Marshello, A. F. J. (1975) Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology*, 32(5), 922-931.
- Kuo, W. J., Sjoström, T., Chen, Y. P., Wang, Y. H. & Huang, C. Y. (2009) Intuition and deliberation: two systems for strategizing in the brain. *Science*, 324(5926), 519-522.
- Kurzban, R. & Leary, D. (2005) Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 102(5), 1803-1807.
- Miller, E. K. & Cohen, J. D. (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167-202.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. J. & Dolan, R. J. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304(5669), 452-454.
- Packard, M. G. & Knowlton, B. J. (2002) Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, 25, 563-593.
- Pruitt, D. G. & Kimmel, M. J. (1977) Twenty years of experimental gaming: critique, synthesis, and suggestions for the future. *Annu Rev Psychol*, 28(1), 363-392.
- Riolo, R. L., Cohen, M. D. & Axelrod, R. (2001) Evolution of cooperation without reciprocity. *Nature*, 414(6862), 441-443.
- Saxe, R. & Wexler, A. (2005) Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, 43(10), 1391-1399.
- Simpson, B. (2004) Social values, subjective transformations, and cooperation in social dilemmas. *Social Psychology Quarterly*, 67(4), 385-395.
- Simpson, B. & Willer, R. (2008) Altruism and indirect reciprocity: the interaction of person and situation in prosocial behavior. *Social Psychology Quarterly*, 71(1), 37-52.

- Smith-Collins, A. P. R., Fiorentini, C., Kessler, E., Boyd, H., Roberts, F. & Skuse, D. H. (2013) Specific neural correlates of successful learning and adaptation during social exchanges. *Social Cognitive and Affective Neuroscience*, 8(8), 887-896.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002) Automated anatomical labeling of activations in Spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273-289.
- Van Lange, P. A. M. (2000) Beyond self-interest: a set of propositions relevant to interpersonal orientations. *European Review of Social Psychology*, 11(1), 297-331.
- Van Lange, P. A. M., Otten, W., De Bruin, E. M. N. & Joireman, J. A. (1997) Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *Journal of Personality and Social Psychology*, 73(4), 733-746.
- Waegeman, A., Declerck, C. H., Boone, C., Seurinck, R. & Parizel, P. M. (2014) Individual differences in behavioral flexibility in a probabilistic reversal learning task: an fMRI study. *Journal of Neuroscience Psychology and Economics*, 7(4), 203-218.
- Watanabe, T., Takezawa, M., Nakawake, Y., Kunimatsu, A., Yamasue, H., Nakamura, M., et al. (2014) Two distinct neural mechanisms underlying indirect reciprocity. *Proceedings of the National Academy of Sciences*.

Supplementary Material

To model learning by experience and to test if experience-based learning affects which neural correlates are activated when players are playing a repeated sequential prisoner's dilemma game, the experiment made use of the experience weighted attraction model (EWA) developed by Camerer and Hua Ho (1999).

This model computes an attraction for each strategy possible in the game. In a prisoner's dilemma, the two possible strategies are to cooperate (C) or to defect (D). In each subsequent trial, the attractions towards these strategies are updated.

The full equation for the attraction, A_C , towards the cooperative strategy, s_C , in trial t is given by:

$$\text{eq. 1} \quad A_C(t) = \frac{\varphi N(t-1)A_C(t-1) + [\delta + (1-\delta) * I(s_C, s(t))] * \pi(s_C, z(t))}{\rho N(t-1) + 1}$$

The first term in the numerator captures what the participant learned from experience in past encounters:

$$\text{eq. 2} \quad \varphi N(t-1)A_C(t-1)$$

where $N(t-1)$ is the number of "observation-equivalents" referring to past encounters, and $A_C(t-1)$ is the attraction towards the cooperative strategy in round $t-1$. The factor φ is a discount factor that weights prior experience in earlier trials less than the experience in the most recent trials. This term captures how strongly participants update their preference towards cooperation, based on built-up experience with this strategy, and therefore captures the learning process of the participants.

The second term describes the influence of the pay-off on the attraction in the current trial t :

$$\text{eq. 3} \quad [\delta + (1-\delta) * I(s_C, s(t))] * \pi(s_C, z(t))$$

The term $\pi(s_C, z(t))$ is the pay-off received when the participant chooses the cooperative strategy, given the response $z(t)$ by the opponent (which can be cooperate or defect). The term $I(s_C, s(t))$ is equal to 1 if the participant chooses to cooperate, and 0 otherwise. Thus, if the participant

cooperates, the pay-off term is equal to the actual pay-off and weighted by 1. If the participant defects, attraction to cooperate is affected by the pay-off that was forgone by not cooperating, weighted by δ .

The sum of these two terms is then normalized by dividing the number of previous encounters (the “observation-equivalents”, $N(t-1)$), discounted by the factor ρ .

The probability of cooperation in trial $t+1$ is defined by the attraction to cooperation vis-à-vis the alternative strategy of defect, A_D :

eq. 4
$$P_C(t + 1) = \frac{e^{\lambda A_C(t)}}{e^{\lambda A_C(t)} + e^{\lambda A_D(t)}}$$

in which λ is a scaling factor. To determine the values of the free parameters, we fitted equation 4 to the data using matlab code by den Ouden *et al.* (2005).

A linear regression model with random effects to account for the differences among participants was estimated, with the estimated probability of cooperation as dependent variable, and trial numbers and SVO as independent variables. This revealed that there is a significant effect of trial number ($B = 0.010$, 95% CI [0.008 0.011], $p < 0.000$, Wald $\chi^2 = 253.63$) but not of SVO. Similar to the results reported in the paper, the interaction between trial number and SVO is also significant ($B = 0.0006$, 95% CI [0.003 0.008], $p = 0.005$, Wald $\chi^2 = 275.00$).

Camerer, C. & Hua Ho, T. (1999) Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4), 827-874.

den Ouden, H. E. M., Frith, U., Frith, C. & Blakemore, S. J. (2005) Thinking about intentions. *Neuroimage*, 28(4), 787-796.

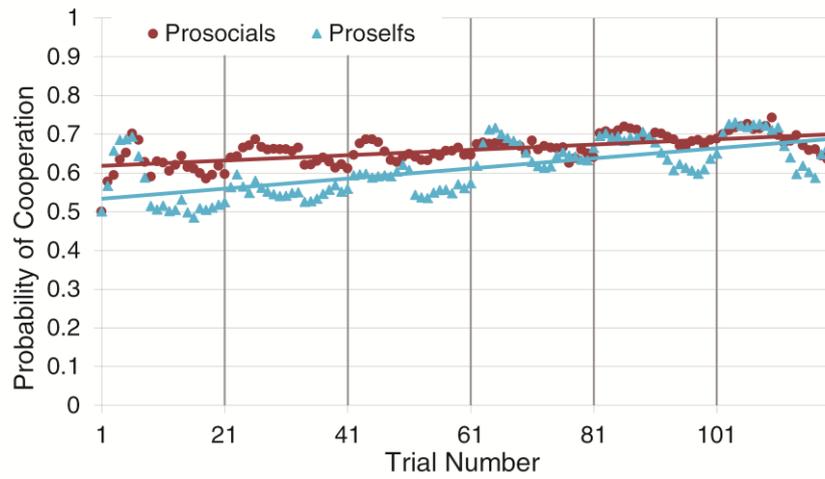


Fig. S1. Estimated probability of cooperative decisions made by proselves or prosocials for each of the 120 trials. Best fit lines, based on least squares, are shown.