Efficient learning mechanisms hold in the social domain and are implemented in the medial prefrontal cortex

Seid-Fatemi, Azade; Tobler, Philippe N

Abstract: When we are learning to associate novel cues with outcomes, learning is more efficient if we take advantage of previously learned associations and thereby avoid redundant learning. The blocking effect represents this sort of efficiency mechanism and refers to the phenomenon in which a novel stimulus is blocked from learning when it is associated with a fully predicted outcome. Although there is sufficient evidence that this effect manifests itself when individuals learn about their own rewards, it remains unclear whether it also does when they learn about others' rewards. We employed behavioral and neuroimaging methods to address this question. We demonstrate that blocking does indeed occur in the social domain and it does so to a similar degree as observed in the individual domain. On the neural level, activations in the medial prefrontal cortex (mPFC) show a specific contribution to blocking and learning-related prediction errors in the social domain. These findings suggest that the efficiency principle that applies to reward learning in the individual domain also applies to that in the social domain, with the mPFC playing a central role in implementing it.

Efficient learning mechanisms hold in the social domain and are implemented in the medial prefrontal cortex

Azade Seid-Fatemi and Philippe N. Tobler

University of Zurich

Number of words: 5,958

Author Note: Azade Seid-Fatemi, Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland; Philippe N. Tobler, Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland.

Correspondence should be addressed to: Azade Seid-Fatemi or Philippe N. Tobler, Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Blümlisalpstrasse 10, CH-8006 Zurich, Switzerland, +41 44 634 50 94, azade.seid-fatemi@econ.uzh.ch or +41 44 634 37 45, phil.tobler@econ.uzh.ch.

Conflict of interest: The authors declare no competing financial interests.

Abstract

When we are learning to associate novel cues with outcomes, learning is more efficient if we take advantage of previously learned associations and thereby avoid redundant learning. The *blocking effect* represents this sort of efficiency mechanism and refers to the phenomenon in which a novel stimulus is blocked from learning when it is associated with a fully predicted outcome. While there is sufficient evidence that this effect manifests itself when individuals learn about their own rewards, it remains unclear whether it also does when they learn about others' rewards. We employed behavioral and neuroimaging methods to address this question. We demonstrate that blocking does indeed occur in the social domain and it does so to a similar degree as observed in the individual domain. On the neural level, activations in the medial prefrontal cortex (mPFC) show a specific contribution to blocking and learning-related prediction errors in the social domain. These findings suggest that the efficiency principle that applies to reward learning in the individual domain also applies to that in the social domain, with the mPFC playing a central role in implementing it.

Keywords: learning theory, social neuroscience, reward, cue competition, medial prefrontal cortex

Introduction

Learning to use environmental cues to predict upcoming events is crucial for adaptive behavior and reward-directed learning. It enables rapid response preparation and execution, thereby potentially providing an advantage over competitors. However, it also requires the use of limited cognitive resources that would be conserved if we prevented its occurrence whenever little or no new information was available. Thus, it would be more efficient if learning only took place when necessary, that is, when previous learning did not make it superfluous. For example, it is efficient not to learn a second cause when a first cause already fully explains an outcome. Associative learning research terms this the *blocking effect* because learning about the previously encountered stimulus "blocks" learning about a concurrently appearing novel stimulus. The underlying rationale is that learning to predict an outcome on the basis of a novel stimulus is redundant if we can already do so on the basis of an antecedent stimulus. By contrast, if the novel stimulus appears concurrently with a stimulus that does not predict the reward, then learning about the novel stimulus is not redundant and blocking does not occur. The blocking effect is expressed in a diminished behavioral response to novel stimuli that provide only redundant information about reward occurrence as compared to novel stimuli that provide non-redundant information. In this sense, the blocking effect represents efficient learning by definition. It should be noted, however, that this does not imply that blocking is also adaptive by definition (indeed coding of redundant information could be adaptive, e.g., when the system is irreducibly noisy). Efficient blocking of novel learning has been

observed in different species including rats (Kamin, 1969), monkeys (Waelti et al., 2001), and humans (Eippert et al., 2012; Prados, 2011; Tobler et al., 2006).

The blocking effect has been investigated primarily in individuals learning about rewards for themselves and has been captured by formal models of learning in the individual domain (Rescorla and Wagner, 1972). However, sometimes rewards are not received by us, but by others. Cues that are redundant for others might nevertheless be of value to us as they may provide us with relevant information that is either relevant now or will become relevant at a later point in time. In particular, these cues may enter different associations with rewards for ourselves as compared to rewards for others. Therefore, in a social context, we might want to keep track of cues concerning rewards for others, even if they are redundant for others. By extension, whether we still rely on previous learning (resulting in blocking) or track all information available to us in the environment (not resulting in blocking) remains an open question for reward learning in the social domain. Indeed, at least with some forms of social learning, such as socially transmitted food preferences, there appears to be little blocking (Galef and Durlach, 1993). In this study, we investigate whether and, if so, how the efficiency principle represented by the blocking effect extends to reward learning in the social domain.

On the neural level, the blocking effect has only been investigated during reward learning in the individual domain. It is expressed in a reduced neural response to blocked stimuli as compared to non-blocked (control) stimuli. This effect has been observed in the ventromedial prefrontal cortex (vmPFC) and the striatum using juice rewards (Tobler et al., 2006) as well as in the amygdala using electric shocks (Eippert et al., 2012). Due to their potential usefulness to oneself, learning about rewards received by others may to a

4

certain extent engage the same neural regions that process rewards received by oneself. Previous work suggests that, in particular, ventral and medial regions of PFC are engaged during reward learning in a social context (Behrens et al., 2008; Burke et al., 2010; Suzuki et al., 2012; Zhu et al., 2012). Thus, if the blocking effect indeed occurred in the social domain as well, the question arises about whether it would be implemented by the same vmPFC regions that implement it in the individual domain. Alternatively, specific mPFC regions may be involved in blocking redundant reward learning in the social domain. To investigate these questions, we developed a social variant of the blocking paradigm and used functional magnetic resonance imaging (fMRI) to examine the role of the mPFC in learning to predict monetary rewards.

Materials and Methods

Participants

Thirty-eight participants (17 female; aged 21.0 ± 0.4 years; range: 18-28) took part in this study. None of the participants had prior histories of neurological or psychiatric disorders and all had normal or corrected-to-normal vision. Written informed consent was obtained from all participants, and the study was approved by the Research Ethics Committee of the Canton of Zurich.

Experimental Design

We investigated the blocking effect (Kamin, 1969) with respect to monetary rewards in the social and individual domains in two separate conditions. Individual rewards were

received by the participant, social rewards by another person. Two female volunteers served as the other person in the social condition. We used two existing persons as volunteers who received money at the end of the experiment to ensure that the consequences of the social rewards were as real as those of the individual rewards. Moreover, using two rather than only one volunteer confederate served to make the social conditions of the experiment more engaging (i.e., more varied and less monotonous) and thereby prevent adaptation. The participants never met the two volunteers face-to-face, but read a brief description of them before the experiment began, which included their initials and information about their gender and age. Before the experiment, we determined the amount of the individual reward according to the social preferences of each subject. We did this to ensure that the rewards in the social and the individual condition had the same subjective value. To achieve this, we used a variant of the Becker–DeGroot–Marschak method (Becker et al., 1964). Specifically, before the experiment, we asked subjects to indicate the amount of money (between CHF 1 and 100) that, if delivered to them, was as valuable as delivering CHF 60 to the other person. The amount of CHF 60 was chosen based on pilot studies with a separate set of subjects showing that CHF 60 yielded affordable individual equivalence amounts (CHF 45.80 ± 1.90). The bid was then compared to a random number between 1 and 100 generated by the computer. If the subject's bid was greater than or equal to the number, they received the indicated amount of money. If their bid was lower than the number, they received nothing and the other person received CHF 60. Thus, the procedure provided an incentive-compatible way of obtaining individual reward amounts that corresponded to the value of social reward amounts. The outcome of the procedure had no influence on

the payout or the number of rewards the participants gained from the actual experiment. The bid was obtained before the experiment and used to set the individual reward amount in the experiment such that it had the same value as the social reward amount, given the subject's social preferences.

As previously described (Tobler et al., 2006; Waelti et al., 2001), the within-subject version of the blocking procedure comprised three consecutive phases. In each of these phases, participants were presented with visual stimuli that were associated with the delivery of different social or individual outcomes (Figure 1A). All of the stimuli used were abstract colored shapes presented on a white background and were similar to those used in previous blocking experiments (Tobler et al., 2006; Waelti et al., 2001). Each trial had either a social or individual outcome, never both, and different stimuli were used for each of the conditions (and thus recipients).

In the first (pretraining) phase, the A (experimental) stimuli (A_{SOCIAL} and $A_{INDIVIDUAL}$) were paired with a social or individual reward. In contrast, the B (control) stimuli (B_{SOCIAL} and $B_{INDIVIDUAL}$) were not paired with a reward. Stimuli were presented 20 times (see Table S1) each and the identities of the stimuli were counterbalanced across participants. Each trial started with a 4-s intertrial interval (ITI) that varied from 2 to 6 s (Figure 1B). Stimuli were presented for 1.5 s at random either to the left or the right of the fixation cross. The outcome was presented concurrently with the stimulus for another 1.5 s. During the presentation of any given stimulus, the participants were to perform a specific key press corresponding to the recipient and to the outcome that would follow the stimulus. In particular, upon each stimulus presentation, participants had to indicate whether they expected reward for self, no reward for self, reward for others, or no reward

7

for others by pressing a key with the index or middle finger of their left or right hand. Thus, there was an individual and social reward key and an individual and social no-reward key and participants were asked to press one of these keys in each trial. This allowed us to measure recipient- as well as outcome-specific learning. Condition-to-hand (individual or social) and key-to-reward (reward or no reward) assignments were counterbalanced across participants. Trials in which the participant failed to respond or responded too late were repeated later. Visual stimuli as well as response recordings were controlled using Cogent 2000 (Wellcome Department of Imaging Neuroscience, London, UK) as implemented in Matlab.

fMRI scanning started in the second (compound conditioning) phase. Visual stimuli were presented on a display that participants viewed via a mirror fitted to the top of the head coil. In the compound phase, A stimuli were presented together with X stimuli (X_{SOCIAL} and $X_{INDIVIDUAL}$), forming rewarded compounds. As a control, B stimuli were presented together with Y stimuli (Y_{SOCIAL} and $Y_{INDIVIDUAL}$) and also followed by a reward. In AX trials, the upcoming reward was predicted by the A stimuli. Therefore, efficiency considerations would suggest that the X stimuli should be blocked from learning. Specifically, in formal learning theory (Rescorla and Wagner, 1972), the blocking effect is explained by "cue competition" and operationalized by summing up the associative strengths (predictive values) of all stimuli present in a given trial. In early AX trials, the sum was already close to the value of the reward itself, leading to little difference between predicted and actual value when the reward occurs. In contrast, in early BY trials, the B stimuli did not predict a reward and the summed associative strength was low. Thus, cue competition explains why Y stimuli, but not X stimuli,

8

should be learned as reward-predicting stimuli. AX and BY trials were presented in 24 trials per condition and intermixed with 14 A and B trials per condition, which served to maintain the previously learned associations (see Table S1). To prevent compound trials in general from being associated with reward, we included control compound trials (CZ trials) that were unrewarded (12 trials each for the social and individual condition; see Table S1). These unrewarded, occasionally interleaved compound stimuli are not standardly used in the blocking procedure. We used them in order to ensure that participants paid attention to each of the individual stimuli that constituted a compound rather than automatically associating the co-occurrence of any two stimuli with reward. We thereby aimed to keep learning more elemental than configural (Melchers et al., 2008).

In a third phase, X and Y stimuli were presented alone in unrewarded test trials. Under the assumption that previous learning blocks subsequent learning, the X stimuli should have been blocked from being associated with social or individual reward, while the Y stimuli should have been associated with reward. Y and X trials were presented in 14 trials each and randomly intermixed with A and B trials (14 trials), AX and BY trials (24 trials), and control compound trials (12 trials), again, to maintain previously learned associations (see Table S1). As before, A, AX, and BY trials were followed by reward in order to maintain the previously learned associations.

During fMRI scanning, the experiment was split into three sessions that did not coincide with conditioning phases in order to prevent rapid extinction of Y stimuli during test trials. The compound conditioning phase spanned the first scanning session and the

first half of the second scanning session. The test phase began with the second half of the second scanning session and ended at the end of the third scanning session.

Participants were instructed that, at the end of the experiment, a portion of the rewards accumulated in correctly predicted trials would be paid out to them and the other two individuals, respectively. To ensure that everyone received approximately the same amount irrespective of their bid in the BDM, we adjusted the percentage for each participant individually. To keep them engaged throughout the task, in each trial in which participants failed to respond or responded too slowly, CHF 1 was deducted from their final monetary payment and the three participants with the highest number of correct responses received an additional payment (CHF 20). The highest percentage of trials missed by a participant was 3%.

Social and individual reward expectation was defined as the percentage of the social and individual reward key pressed, respectively. Reward expectations were evaluated using paired t-tests and two-way repeated-measures ANOVAs. The degree of participant-specific behavioral blocking was calculated as the difference between recipient-specific reward key presses for Y stimuli and those for X stimuli. The larger the difference, the stronger the blocking effect. Comparing the responses to Y with those to X stimuli is the standard approach to determining whether blocking has taken place. However, for the neural data, we also analyzed outcome-related activation in AX and BY trials to measure the differential learning responses when comparatively small or large amounts of learning occur, respectively (see above and below).

fMRI Data Acquisition

Functional magnetic resonance (fMRI) data were acquired on a Philips Achieva 3T whole-body scanner equipped with an eight-channel head coil (Philips Medical Systems, Best, The Netherlands) at the Laboratory for Social and Neural Systems Research, University of Zurich. We acquired gradient-echo T2*-weighted echoplanar images (EPIs) with blood-oxygen-level-dependent (BOLD) contrast (slices/volume, 33; repetition time, 1.75 s). Approximately 530-710 volumes were collected per session (variation was due to experiment phase and individual differences in the number of repeated trials) along with 5 "dummy" volumes at the start of the scanning session to allow for magnetization to stabilize to a steady state. Scan onset times varied relative to stimulus onset times. Slice orientation was tilted 20° away from the anterior commissure-posterior commissure line, caudal > rostral. Imaging parameters were: echo time, 30 ms; field-of-view, 240 mm; in-plane resolution, 3 mm; slice thickness, 3 mm; interslice gap, 0.75 mm. A T1-weighted structural image was also acquired for each participant. These high-resolution T1-weighted structural scans were coregistered to their mean EPIs and averaged to permit anatomical localization of the functional activations at the group level.

fMRI Data Analysis

fMRI data processing and statistical analyses were carried out using statistical parametric mapping (SPM8; Wellcome Department of Imaging Neuroscience, London, UK). Data preprocessing consisted of realignment, coregistration, segmentation, spatial normalization using the DARTEL toolbox and smoothing using a Gaussian kernel with a full width at half maximum of 10 mm. Data analysis was performed using a general linear model (GLM) approach. The first-level design matrix of each participant included

separate regressors for each of the four learned stimulus conditions (A$_{INDIVIDUAL}$, A$_{SOCIAL}$, B$_{INDIVIDUAL}$, B$_{SOCIAL}$) modelled at the event-onset time, the compound conditioning trials at the time of the outcome (AX$_{INDIVIDUAL}$, AX$_{SOCIAL}$, BY$_{INDIVIDUAL}$, BY$_{SOCIAL}$) to capture prediction error-related responses during learning, and the four test trial types modelled at the event-onset time (X$_{INDIVIDUAL}$, X$_{SOCIAL}$, Y$_{INDIVIDUAL}$, Y$_{SOCIAL}$) to capture blocking. In order to identify brain regions that correlate with prediction error during compound learning trials, we parametrically modulated the AX and BY regressors with trial-wise and mean-corrected prediction errors (δ) derived from a standard reinforcement learning model (see below). To account for the variance that can be explained by stimulus presentation, we created two additional regressors for compound conditioning trials at event-onset time, combining AX and BY trials into a single regressor (AX/BY$_{SOCIAL}$, AX/BY$_{INDIVIDUAL}$). Finally, we included regressors of no interest for the unrewarded compound trials and for participant-specific movement parameters (3 regressors for rotation and 3 for translation). All regressors were convolved with the canonical hemodynamic response function (HRF). For each regressor, we included all trials irrespective of the participants' response.

In order to identify brain regions involved in prediction error-based learning during compound conditioning, we parametrically modulated the AX and BY regressors with mean-corrected prediction errors derived from a variant of a simple reinforcement learning model (Rescorla and Wagner, 1972). In each trial, prediction errors were computed according to $\delta_t = \alpha (\lambda_t - V_t)$, where V_t corresponds to the value V predicted by all stimuli presented in trial t, λ_t corresponds to the reward in trial t, and α corresponds to the learning rate. The learning rate determines how much weight is given to recent

12

experience as captured by the prediction error. It is a free parameter that can be used to characterize how quickly participants learn in different conditions (see e.g. Burke et al., 2010). We estimated the learning rate by fitting the prediction error model above to the trial-by-trial percentage of reward keypress responses in BY trials, averaged across participants. These keypresses are a measure of participants' reward prediction in a given trial. As different keys were used for the prediction of social and individual reward, we were able to estimate separate learning rates by using their keypresses for the social and individual condition, respectively. The estimated learning rate was 0.10 for the social condition and 0.15 for the individual condition (no significant difference).

The prediction error in a given trial is used to update the associative strengths of all stimuli present in that trial. For example, in the initial BY trials, the associative strength of BY is low, so a reward should generate a positive prediction error. During training with the BY compound, the reward becomes more and more predictable and, according to theory, the prediction error gradually decreases (Fig. S1). On the other hand, in the initial AX trials, the associative strength of AX is already high (i.e., reward is already fully predicted) due to the pretrained A stimuli, so a decrease in prediction error should not occur. Taken together, for regions involved in learning, over the course of compound conditioning, we expect to observe a greater reduction in prediction-error-related activity in BY than in AX trials. This would be captured by a better fit with a parametric modulator that models a decreasing prediction error signal in BY compared to AX trials.

Linear contrasts of regression coefficients of A vs. B (stimulus response), BY vs. AX (prediction error modulator), and Y vs. X (stimulus response) were computed at the

13

single-participant level and then taken to group-level analyses where we used one-sample t-tests or correlations with participant-specific degree of blocking in the social or individual domain. Correction for multiple comparisons (FWE, $p < 0.05$) was performed either in areas of interest or at the whole-brain level. Our a priori region of interest for the individual condition was defined functionally as a 15-mm sphere around the peak of a previously reported coordinate reflecting individual blocking in vmPFC (-18, 36, -6) (Tobler et al., 2006). Moreover, we assessed blocking effects within spheres around peak activations identified by the independent reward expectation contrast of A vs. B. Outside these regions of interest, correction for multiple comparisons was performed at the whole-brain cluster level ($p < 0.05$, cluster-inducing threshold: $p < 0.001$). In the figures, the left side of the brain is shown on the left.

Results

Behavioral Results

We used a within-subject design that employed three phases to test blocking in the social and individual domain (Figures 1A and 1B, see Methods). In the pre-training phase, participants learned to associate one stimulus (A) with reward and another (B) with no reward; they maintained these associations in the subsequent compound and test phases. The association between A and reward was expected to block individual learning in the compound phase, but it was an open question whether this would also hold for social learning. The degree to which blocking had occurred was assessed in unrewarded trials in the test phase. First, we tested whether participants had learned the (previous) stimulus-

outcome associations: Participants predicted recipient-specific reward outcomes when presented with reward-predicting A stimuli, but not when presented with B (control) stimuli (Figure 1C). This resulted in a significantly higher number of reward key presses for A vs. B stimuli (over all phases), for both social and individual conditions (social: 92.0% ± 1.3% vs. 14.1% ± 2.4%, $t_{(37)} = 23.75$; individual: 96.5% ± 0.7% vs. 5.3% ± 1.4% (mean ± SEM), $t_{(37)} = 47.34$, both $p < 0.001$). The results were very similar when the analysis was limited to A and B trials of the pretraining phase (social: 88.2% ± 2.0% vs. 17.0% ± 3.3%, $t_{(37)} = 15.08$; individual: 95.1% ± 1.3% vs. 6.9% ± 2.4% (mean ± SEM), $t_{(37)} = 26.23$, both $p < 0.001$), indicating that the pretraining phase, which took place before the compound conditioning phase, was successful in both the social and the individual condition. Taken together, the participants learned to discriminate in an outcome- and recipient-specific manner between stimuli predicting reward and stimuli predicting no reward.

Next, we investigated whether the course of learning differed in the BY vs. AX condition. In the BY condition, reward key presses gradually increased in both the social and the individual condition (Figure 1D). Thus, in early trials, participants showed only a low number of reward key presses, but learned the association over time. In contrast, there was only a very mild increase in reward key presses in the AX condition as participants were already predicting the reward outcome for AX at the beginning of the compound phase. To assess whether the increase in reward key presses in BY trials differed from that in AX trials, we compared the first six trials (early) with the last six trials (late) and found that, in both the social and the individual condition, trial type (BY vs. AX) interacted with time (early vs. late compound trials; social: $F_{(1,37)} = 54.82$;

15

individual: $F_{(1,37)} = 257.34$; both $p < 0.001$), indicating that more learning occurred during BY trials than during AX trials.

We then assessed whether the blocking effect manifests itself not only in the individual but also in the social domain by comparing participants' reward expectation to potentially blocked X stimuli vs. Y (control) stimuli in non-rewarded test trials. There was an increase in reward key presses for Y (control) stimuli as compared to X stimuli in both the social ($t_{(37)} = 3.07$, $p < 0.005$) and the individual ($t_{(37)} = 2.48$, $p < 0.05$) condition (Figure 1C). Although the difference between Y and X was smaller than that between A and B (social: $t_{(37)} = 8.76$; individual: $t_{(37)} = 11.06$, both $p < 0.001$), participants clearly treated Y and X differently in the individual as well as the social condition. If participants had failed to learn anything about the outcome of a given stimulus, we would expect performance at chance level, as there was a 50% chance of pressing either the reward or the no-reward key. For the social as well as the individual condition, reward key presses for X stimuli did not differ from chance (social: $51.5\% \pm 6.8\%$, $t_{(37)} = 0.22$, $p = 0.83$; individual: $52.8\% \pm 6.8\%$, $t_{(37)} = 0.41$, $p = 0.68$) whereas those for Y (control) stimuli occurred more often than 50% (social: $69.7\% \pm 5.8\%$, $t_{(37)} = 3.40$; individual: $68.4\% \pm 6.2\%$, $t_{(37)} = 2.97$, both $p < 0.05$), again confirming that blocking had occurred in the individual as well as in the social condition.

To measure participant-specific differences in the blocking effect and compare social and individual blocking, we determined each participant's degree of blocking by calculating the difference between reward key presses for Y and X. Interestingly, the degree of blocking was similar for the social and individual condition ($t_{(37)} = 0.53$, $p = 0.6$). Moreover, across participants, the degree of blocking in the social condition was

correlated with the degree of blocking in the individual condition ($R^2 = 0.45$, $p < 0.001$). Thus, from a behavioral standpoint, blocking in the social and individual conditions were related.

We also tested whether the social and individual conditions differed in their salience as assessed by differences between the conditions with respect to response time. There were no significant differences between social and individual test trials with respect to response times (X: 853.8 ± 14.9 ms vs. 852.6 ± 15.6, $t_{(37)} = 0.10$, $p = 0.92$; Y: 823.0 ± 16.2 ms vs. 839.4 ± 18.6, $t_{(37)} = -0.81$, $p = 0.42$). In A and B trials, participants responded faster in the individual as compared to the social condition (A: 729.2 ± 9.0 ms vs. 785.3 ± 11.5 ms, $t_{(37)} = -6.04$; B: 748.0 ± 10.6 ms vs. 790.8 ± 10.5 ms, $t_{(37)} = -5.50$; both $p < 0.001$). Thus, while individual A and B trials may have been more salient than social A and B trials, there is no evidence for a difference in salience between social and individual Y and X trials, with which blocking was assessed at the neural level.

fMRI Results

Blocking in the social domain

First we investigated responses in A vs. B trials with respect to the participants' expectation that another person would receive a monetary reward (social condition). We found stronger activation in the mPFC when participants expected another person to receive a reward than when they did not (Figure 2A; 2, 60, 16; $t_{(37)} = 5.41$; $p < 0.05$, whole-brain FWE cluster-level corrected; see Table 1 for additional whole-brain corrected activations).

To test whether the mPFC activity found for A vs. B shows a blocking effect as well, we investigated the contrast of Y vs. X. This is the standard contrast (Eippert et al., 2012; Tobler et al., 2006) used to test for the blocking effect as it captures reduced neural responses to the (blocked) X stimulus as compared to the (non-blocked) Y stimulus. An increased response to the Y compared to the X stimulus reflects the stronger reward prediction for Y vs. X, similar to the difference in reward prediction for A vs. B, but uncontaminated by actual reward delivery. We therefore used a mask including a 10-mm sphere around the peak coordinate from the contrast of A vs. B in the social condition and performed an independent second-level correlation analysis of differential brain activation in Y vs. X against the participant-specific behavioral difference of Y vs. X. We found a correlation in mPFC (Figures 2B and 2C; 6, 60, 12; $t_{(36)} = 3.63$; $p < 0.05$, FWE small-volume corrected). Thus, within mPFC, similar subregions showed activations reflecting reward expectation and blocking in the social domain.

Blocking in the individual domain

To confirm previous findings on blocking in the individual domain, we examined responses related to reward expectation and blocking in the individual condition. Specifically, we analyzed activity in the vmPFC, a region identified in a previous study on the blocking effect involving liquid reward (Tobler et al., 2006). We found that activity in the vmPFC was stronger for reward-predicting stimuli than for neutral stimuli (Figure S2A; -4, 40, -6; $t_{(37)} = 3.81$; $p < 0.05$, FWE small-volume corrected) and increased with the degree of behavioral blocking (Figures S2B and S2C: -6, 42, -4;

$t_{(36)} = 3.84$; $p < 0.05$, FWE small-volume corrected). These data suggest that vmPFC activations reflect both blocking and reward expectation in the individual condition.

Comparison of social and individual conditions

On the behavioral level, we found similar reward expectation and blocking effects for the social and individual conditions. Nevertheless, it is possible that distinct regions in the brain keep track of the reward recipient. We therefore tested whether the mPFC response reflecting blocking and reward expectation is stronger in the social than in the individual domain. First, we assessed whether activation for A vs. B is specific for the social condition. The more dorsal part of the mPFC that was identified for social learning and blocking (Figures 2A and 2B) responded more strongly in the social than in the individual condition (A vs. B social > A vs. B individual: 8, 56, 12; $t_{(37)} = 4.52$; Figures 3A and 3B; $p < 0.05$, whole-brain FWE cluster corrected; additional whole-brain corrected activations are shown in Table 1). Moreover, activity in the same region was also stronger for blocking in the social than in the individual condition. In other words, we found activity in mPFC for the contrast of Y vs. X social > Y vs. X individual (12, 56, 10; $t_{(37)} = 3.57$; Figures 3C and 3D; $p < 0.05$, FWE small-volume corrected in a 10-mm sphere around the peak coordinate of A vs. B social > A vs. B individual). This activation extended into more lateral parts of the mPFC with an additional peak at 26, 60, 6 ($p < 0.05$, whole-brain FWE cluster-level corrected). Note that these preferential neural effects of blocking in the social condition occurred in the absence of significant behavioral or value differences between the individual and the social conditions.

Development of blocking in the social domain

After establishing that the mPFC plays a preferential role in blocking in the social domain, we investigated social learning during the compound conditioning phase. In the AX trials, the social reward was already fully predicted by the pretrained stimulus A and therefore the reward was expected to elicit little or no prediction error signal. In contrast, in the BY trials, social reward was not predicted as B had not been rewarded in pretraining. Consequently, the reward outcome was expected to generate a sizeable prediction error in early BY trials. As learning progressed from trial to trial, the reward outcome was expected to elicit a gradually decreasing prediction error. Accordingly, we tested for better fits with decreases in prediction error in BY compared to AX trials. We therefore parametrically modulated the AX and BY regressors in the social condition with trial-wise mean-corrected prediction errors derived from a simple reinforcement learning model (see Methods and Figure S1). We found that activation in the dorsomedial PFC (dmPFC) fitted the parametric modulator for BY compared to AX trials better, reflecting the more strongly decreasing prediction error responses in BY trials in the social condition (Figures 4A and 4B; 10, 30, 38; $t_{(37)} = 4.65$; $p < 0.05$, FWE cluster corrected; see Table 1 for additional whole-brain corrected activation). Additionally, at less stringent statistical thresholds, we found that the differential dmPFC activation was more sensitive to prediction errors in the social as compared to the individual condition (BY vs. AX social > BY vs. AX individual: 12, 30, 40; $t_{(37)} = 3.22$; $p < 0.001$, uncorrected). Furthermore, the differential fit of the prediction-error-related activity in the dmPFC correlated with the degree of behavioral blocking in the social condition (-12, 26, 40; $t_{(36)} = 3.40$; $p < 0.001$, uncorrected). Although both findings should be interpreted

20

with care due to their uncorrected nature, they are in line with the notion that the dmPFC preferentially codes prediction errors during reward learning in the social rather than the individual domain and that this prediction error coding is related to participant-specific differences in blocking in the social domain.

Discussion

In this study, we investigated whether and, if so, how the efficiency principle represented by the blocking effect extends to reward learning in the social domain. Our behavioral results did indeed reveal a blocking effect in the social domain and thereby suggest that, as in the individual domain, efficiency is weighted more heavily than the complete encoding of all available information. Thus, the same mechanism that leads to efficient reward learning in the individual domain also serves as an efficient strategy to optimize learning in the social domain. Moreover, the degree to which the effect manifested itself in the two domains was correlated across participants. Nevertheless, although we found similar and correlated blocking effects in the two domains on the behavioral level, on the neural level, we found that the more dorsal mPFC assumes a preferential role for the blocking of socially relevant cues.

At the behavioral level, we found blocking not only with individual, but also with social learning. Thus, our study suggests that blocking occurs in at least some forms of social learning. This was not obvious from the outset as attempts to show blocking, for example, in the domain of socially transmitted food preferences were not successful (Galef and Durlach, 1993). In the case of socially transmitted food preferences, however,

the definition of the unconditioned stimulus (social interaction with a demonstrator rat) and its relationship to the dependent variable (food consumption by the observer rat) is less obvious than in more standard paradigms of individual learning. By contrast, we used similar response requirements and clearly defined rewards in both individual and social conditions, which facilitated the comparison of individual and social learning.

Blocking and reward expectation effects in the social domain were enhanced over and above those in the individual domain in relatively more dorsal regions of mPFC. This preferential relationship with social effects arose even though we equated subjective values and response requirements in the individual and social conditions (see Methods). Thus, we can exclude the possibility that the mPFC activation simply reflects differential values or response requirements related to one's own or others' rewards. It is not likely due to differences in salience either (e.g., Leathers and Olson, 2012) as response times were similar in X and Y trials in the social and individual conditions. These conclusions are further supported by a control analysis: We obtained subjective desirability ratings indicating how much each participant cared about the two other people receiving a monetary payoff during the task. Including this variable as a covariate of no interest allowed us to identify a very similar increase in activation in mPFC that correlated with the degree of blocking observed in the social condition. Thus, the preferential contribution of the more dorsal mPFC to social learning and blocking cannot be explained by differential value- and salience-related effects.

Incidentally, the control analysis also renders an explanation in terms of conflict less likely: Those who cared less about others receiving a reward should have felt more conflict when they had to perform a movement that would be followed by such a reward.

22

However, the desirability ratings were not related to participant-specific differences in response times during social reward expectation or blocking. In line with the absence of a role of conflict, our activations occurred in more anterior and ventral locations of mPFC than those typically associated with conflict (Botvinick et al., 2001; Kerns et al., 2004; Shenhav et al., 2013).

Our data suggest that the relatively dorsal mPFC regions contribute to other-directed reward learning by implementing an efficient learning mechanism originally described in empirical studies and formal models of individual learning (Kamin, 1969; Rescorla and Wagner, 1972). Our data thereby converge with reports of relatively dorsal mPFC involvement in other aspects of social learning. For example, Behrens et al. (2008) found that activity in the dmPFC correlates with errors in the predicted helpfulness of confederate advice. In another case, dmPFC activation during an inspection game was found to correlate with the degree to which players thought they influenced their opponent's behavior (Hampton et al., 2008). Thus, activity of the dorsal mPFC can be captured particularly well with formal models of social learning with the unifying explanation that this region encodes social reward prediction errors.

Responses in the dmPFC reflected the gradual decrease in prediction errors in BY trials, indicating that this region processes the change in prediction errors as anticipated by formal learning theory. Note that this dmPFC activation is more dorsal and posterior than the mPFC region we found to be sensitive to blocking in the social domain, suggesting that different subregions of the dmPFC are engaged at different stages of social learning. Future research may therefore focus on the mechanisms underlying the development of blocking in the social domain and investigate in more detail how the

development of the effect in the compound phase relates to its expression in the test phase.

In our social condition, participants predicted whether another person would obtain a reward. Outcomes related to others may be more abstract than one's own outcomes (Amodio and Frith, 2006). In this sense, the present findings support the idea of a dorsal-ventral and posterior-anterior axis (Denny et al., 2012; Koritzky et al., 2013; Suzuki et al., 2012) according to which the more dorsal and anterior mPFC processes more abstract and complex information than the more ventral and posterior mPFC. This in turn is in agreement with the core role of anterior mPFC in social value processing, other-related judgments, and mentalizing (Amodio and Frith, 2006; Fareri et al., 2012; Gilbert et al., 2006; Krienen et al., 2010; Mitchell, 2009; Ochsner et al., 2004). Indeed, the most anterior part of the prefrontal cortex, the frontal pole, may have emerged as a new prefrontal area during primate evolution (Genovesio et al., 2014). Together with the notion that social functions developed to a disproportionate degree in the later stages of primate evolution (Dunbar, 1998), it is tempting to speculate that this area might have evolved to serve a preferential role for learning about observed and socially relevant outcomes.

In the domain of causal learning, the blocking effect also occurs in an observational context in which the participant has to learn causal relationships between actions or events and their associated outcomes (Dickinson et al., 1984). We cannot rule out the possibility that the neural results obtained in our study might also generalize to blocking effects in causal learning and may be partly driven by explicit (verbal) reasoning. Previous studies primarily found the lateral PFC to be crucial for causal

learning and blocking of causal learning (Corlett and Fletcher, 2012; Fletcher et al., 2001; Turner et al., 2004) and verbal reasoning (Costafreda et al., 2006; Tsuchida and Fellows, 2013). Extending these studies, we found a more medial region that was specifically involved in blocking in the social over and above the individual domain.

Although the present study focused on blocking in the social domain, we also found individual blocking effects. These were represented primarily in the ventral part of the mPFC and, at less stringent statistical thresholds, also in the striatum and the posterior cingulate (data not shown). These findings replicate our and others' previous reports on individual blocking and learning (McDannald et al., 2014; Tobler et al., 2006; note that in some of the previous studies subjects were thirsty and rewards were drops of liquid, which could have resulted in a more homogeneous and higher value of the reward). By using secondary (monetary) rather than primary rewards to study the neural basis of the blocking effect, we go beyond previous research and show that the vmPFC contributes to individual neural blocking, not only in the context of primary rewards (liquid), but also in that of secondary rewards (money). It should be noted, however, that we did not find significantly stronger activation in the vmPFC for the direct comparison between the individual and social condition. Thus, we cannot conclude that the ventral part of mPFC is specific for processing self-relevant rewards. Indeed, there have also been reports of other-relevant learning processes in the vmPFC (Burke et al., 2010; Suzuki et al., 2012). One possibility worthy of further study is that these ventral regions are engaged when other-relevant learning has a direct benefit (instrumental value) for the observing individual (Burke et al., 2010). In contrast, the observation of others' rewards had comparatively little instrumental value in our study. Thus, it remains to be determined

what specific contextual aspects lead to vmPFC contribution during learning in the social domain.

Taken together, our findings substantiate the notion that the same formal learning processes hold and facilitate efficient learning in both the individual and the social domain. Moreover, our data indicate that regions of dorsal mPFC play a preferential role for implementing these processes when rewards are socially relevant.

References

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Review Neuroscience*, 7, 268-277.

Becker, G.M., Degroot, M.H., Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9, 226-232.

Behrens, T.E.J., Hunt, L.T., Woolrich, M.W., Rushworth, M.F.S. (2008). Associative learning of social value. *Nature*, 456, 245-249.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624-652.

Burke, C.J., Tobler, P.N., Baddeley, M., Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences USA*, 107, 14431-14436.

Corlett, P.R., Fletcher, P.C. (2012). The neurobiology of schizotypy: Fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. *Neuropsychologia*, 50, 3612-3620.

Costafreda, S.G., Fu, C.H.Y., Lee, L., Everitt, B., Brammer, M.J., David, A.S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. *Human Brain Mapping*, 27, 799-810.

Denny, B.T., Kober, H., Wager, T.D., Ochsner, K.N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, 24, 1742-1752.

Dickinson, A., Shanks, D., Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology A*, 36, 29-50.

Dunbar, R.I.M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178-190.

Eippert, F., Gamer, M., Büchel, C. (2012). Neurobiological mechanisms underlying the blocking effect in aversive learning. *Journal of Neuroscience*, 32, 13164-13176.

Fareri, D.S., Niznikiewicz, M.A., Lee, V.K., Delgado, M.R. (2012). Social network modulation of reward-related signals. *Journal of Neuroscience*, 32, 9045-9052.

Fletcher, P.C., Anderson, J.M., Shanks, D.R., Honey, R., Carpenter, T.A., Donovan, T., Papadakis, N., Bullmore, E.T. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nature Neuroscience*, 4, 1043-1048.

Galef, B.G., Durlach, P.J. (1993). Absence of blocking, overshadowing, and latent inhibition in social enhancement of food preferences. *Animal Learning and Behavior*, 21, 214-220.

Genovesio, A., Wise, S.P., Passingham, R.E. (2014). Prefrontal-parietal function: From foraging to foresight. *Trends in Cognitive Sciences*, 18, 72-81.

Gilbert, S.J., Spengler, S., Simons, J.S., Steele, J.D., Lawrie, S.M., Frith, C.D., Burgess, P.W. (2006). Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *Journal of Cognitive Neuroscience*, 18, 932-948.

Hampton, A.N., Bossaerts, P., O'Doherty, J.P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences USA*, 10, 6741-6746.

Kamin, L. (1969). Predictability, surprise, attention and conditioning. In B.A. Campbell, R.M. Church (eds.), *Punishment and aversive behavior* (pp. 279-296). New York: Appleton-Century-Crofts.

Kerns, J.G., Cohen, J.D., MacDonald III, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303, 1023–1026.

Koritzky, G., He, Q., Xue, G., Wong, S., Xiao, L., Bechara, A. (2013). Processing of time within the prefrontal cortex: Recent time engages posterior areas whereas distant time engages anterior areas. *NeuroImage*, 72, 280-286.

Krienen, F.M., Tu, P.C., Buckner, R.L. (2010). Clan mentality: Evidence that the medial prefrontal cortex responds to close others. *Journal of Neuroscience*, 30, 13906-13915.

Leathers, M.L., Olson, C.R. (2012). In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science*, 338, 132-135.

McDannald, M.A., Jones, J.L., Takahashi, Y., Schoenbaum, G. (2014). Learning theory: A driving force in understanding orbitofrontal function. *Neurobiology of Learning and Memory*, 108, 22-27.

Melchers, K.G., Shanks, D.R., Lachnit, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioural Processes*, 77, 413-427.

Mitchell, J.P. (2009). Social psychology as a natural kind. *Trends in Cognitive Sciences*, 13, 246-251.

Ochsner, K.N., Knierim, K., Ludlow, D.H., Hanelin, J., Ramachandran, T., Glover, G., Mackey, S.C. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, 16, 1746-1772.

Prados, J. (2011). Blocking and overshadowing in human geometry learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 37, 121-126.

Rescorla, R.A. and Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black, W.F. Prokasy (eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Shenhav, A., Botvinick, M.M., Cohen, J.D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, 79, 217-240.

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J., Ichinohe, N., Haruno, M., Cheng, K., Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron*, 74, 1125-1137.

Tobler, P.N., O'Doherty, J.P., Dolan, R.J., Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*, 95, 301-310.

Tsuchida A., Fellows L.K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex*, 49, 1790-1800.

Turner, D.C., Aitken, M.R.F., Shanks, D.R., Sahakian, B.J., Robbins, T.W., Schwarzbauer, C., Fletcher, P.C. (2004). The role of the lateral frontal cortex in causal associative learning: Exploring preventative and super-learning. *Cerebral Cortex*, 14, 872-880.

Waelti, P., Dickinson, A., Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, 412, 43-48.

Zhu, L., Mathewson, K.E., Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proceedings of the National Academy of Sciences USA*, 109, 1419-1424.

Table 1. Brain regions exhibiting additional learning- or blocking-related activation.

Brain region	x	y	z	t
A > B (Social)				
Medial prefrontal cortex	2	60	16	5.41
Posterior cingulate cortex	-6	-52	16	4.74
A > B (Social vs. Individual)				
Rolandic operculum	-60	-6	10	5.32
Precuneus	8	-54	58	4.81
Medial prefrontal cortex	8	56	12	4.52
Middle occipital gyrus	48	-78	20	4.35
Y > X (Social vs. Individual)				
Lateral prefrontal cortex	26	60	6	4.36
BY_PM > AX_PM (Social)				
Parietal cortex	34	62	38	4.84
Dorsomedial prefrontal cortex	10	30	38	4.65

Regions that survive whole-brain FWE correction at the cluster level, with a cluster-inducing threshold of $p < 0.001$. Coordinates are denoted by x, y, z (in mm; MNI space).

Fig. 1 Experimental design and behavioral results. *A*, Three phases of blocking paradigm with monetary rewards. During pretraining, participants learned to associate stimuli with the presence or absence of monetary outcomes. Reward-predicting A stimuli were followed by a monetary reward, but not neutral B stimuli (each of these had a social and an individual variant, see Fig. 1B). During compound conditioning, X and Y stimuli appeared together with A and B stimuli in rewarded compounds. In AX trials, the reward was fully predicted by the A stimulus. Therefore, the X stimulus was expected to be blocked from learning. In contrast, in BY trials, the reward was not predicted by the B stimulus, so the Y (control) stimulus was expected to be learned as a reward-predicting stimulus. During the test phase, X and Y stimuli were presented alone and remained unrewarded. While learned stimulus Y was expected to predict upcoming reward, blocked stimulus X was not. During the compound and the test phase, trial types of the previous phases were also presented in order to maintain learned associations. *B*, Example of pretraining trials. Abstract visual stimuli were presented in random order, either to the left or the right of the fixation cross. Upon presentation of a stimulus, the participants were to perform a specific key press corresponding to the recipient (self or other) and to the outcome (reward or no reward) that would follow the stimulus. The outcome was shown together with the stimulus for another 1.5 s. The ITI varied between 2 and 6 s. *C*, Participants showed an increase in reward-expecting responses (quantified as percentage of key presses) to reward-predicting A stimuli and Y (control) stimuli as compared to unrewarded B stimuli and blocked X stimuli in the social as well as the

individual condition, suggesting blocking effects in both cases. Key presses for A and B are shown for trials from all three phases. Error bars indicate SEMs. *D*, Mean learning curves averaged across participants showed a stronger increase in the BY as compared to the AX condition for the social as well as the individual condition. Shown is the percentage of reward key presses over time (each bin averaged over six trials). Error bars indicate SEMs.

Fig. 2 Activity in mPFC reflects expectation and blocking of social rewards. *A*, mPFC responses were higher to reward-predicting A stimuli as compared to neutral B stimuli (2, 60, 16; $p < 0.05$, FWE-corrected). Contrast estimates (inset) show mPFC responses to A and B stimuli separately. Error bars indicate SEM. *B* and *C*, Differences in activation responses to Y as compared to X increased in the mPFC (6, 60, 12, $p < 0.05$, FWE small-volume corrected) with degree of blocking in the social condition. Blocking was quantified as the difference in reward-expecting responses to nonblocked stimulus Y compared to blocked stimulus X. Color bars indicate z-scores.

Fig. 3 Preferential responses to reward expectation and blocking in the social as compared to the individual condition. *A* and *C*, Stronger responses in the mPFC in the social than in the individual condition for the contrast of A vs. B (8, 56, 12, $p < 0.05$, FWE small-volume corrected) and Y vs. X (12, 56, 10, $p < 0.05$, FWE small-volume corrected). *B* and *D*, Contrast estimates show mPFC response for the contrast A vs. B and Y vs. X separately for the social and the individual condition. Error bars indicate SEM. Color bars indicate z-scores.

Fig. 4 Differential activations during the development of the blocking effect. *A* Activation in the dmPFC (peak at 10, 30, 38; $p < 0.05$, FWE cluster corrected) was better fitted by a parametric modulator that modeled decreases in prediction error in BY (PM_BY) as compared to AX trials (PM_AX). Color bar indicates z-score. *B*, Contrast estimates show response in the dmPFC for the parametric modulators of BY and AX separately. Error bars indicate SEM.
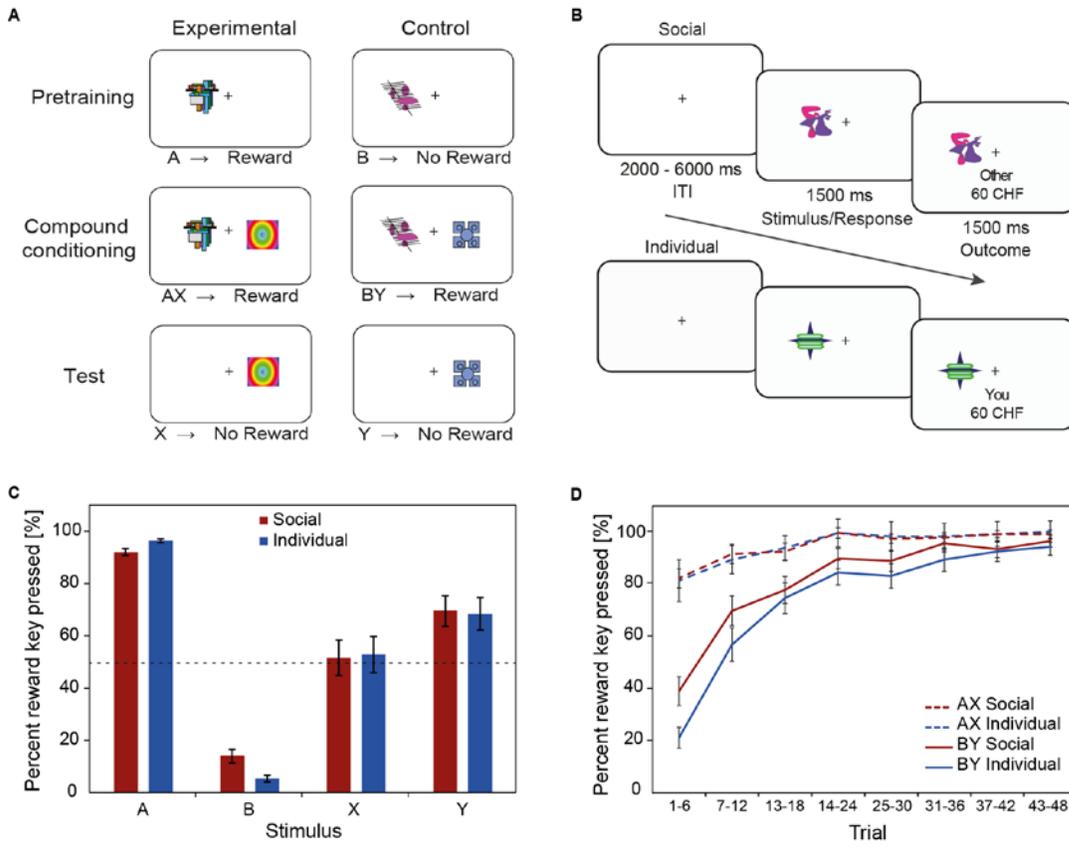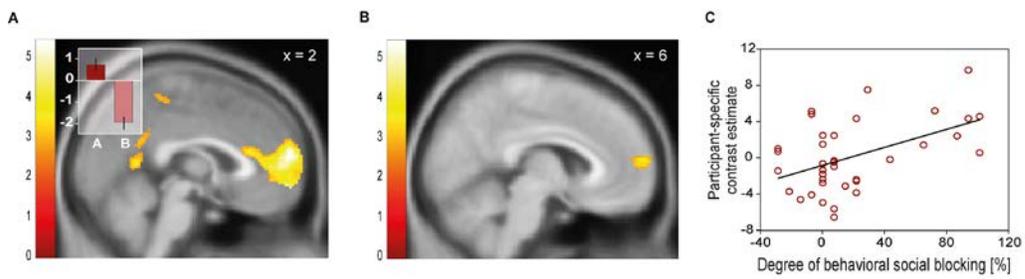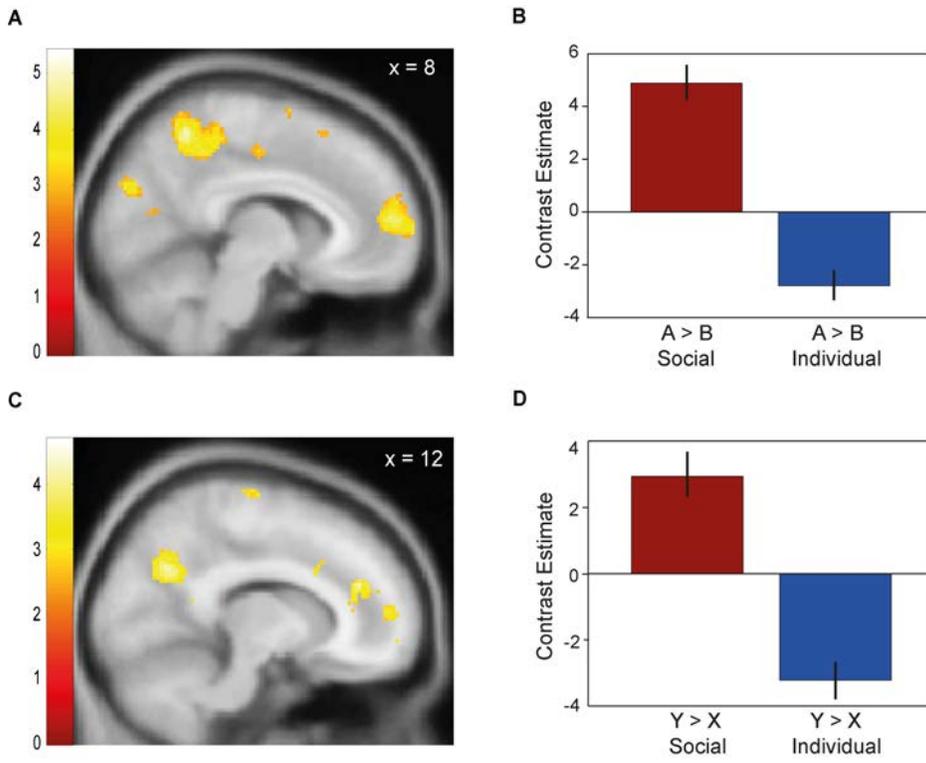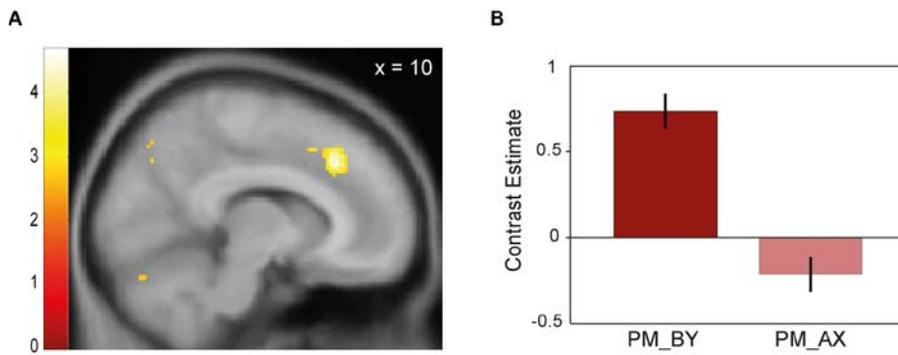
Figure 1

Figure 2

Figure 3

Figure 4

Supplementary data for:

Efficient learning mechanisms hold in the social domain and are implemented in the

medial prefrontal cortex

Azade Seid-Fatemi and Philippe N. Tobler

University of Zurich

Table S1. Trial types and number of trials used in each phase

Phase	Social condition		Individual condition	
	Trial type	**Number of trials**	**Trial type**	**Number of trials**
Pretraining	A_{SOCIAL} +	20	$A_{INDIVIDUAL}$ +	20
	B_{SOCIAL} -	20	$B_{INDIVIDUAL}$ -	20
Compound conditioning	A_{SOCIAL} +	14	$A_{INDIVIDUAL}$ +	14
	B_{SOCIAL} -	14	$B_{INDIVIDUAL}$ -	14
	AX_{SOCIAL} +	24	$AX_{INDIVIDUAL}$ +	24
	BY_{SOCIAL} +	24	$BY_{INDIVIDUAL}$ +	24
	CZ_{SOCIAL} -	12	$CZ_{INDIVIDUAL}$ -	12
Test	A_{SOCIAL} +	14	$A_{INDIVIDUAL}$ +	14
	B_{SOCIAL} -	14	$B_{INDIVIDUAL}$ -	14
	AX_{SOCIAL} +	24	$AX_{INDIVIDUAL}$ +	24
	BY_{SOCIAL} +	24	$BY_{INDIVIDUAL}$ +	24
	CZ_{SOCIAL} -	12	$CZ_{INDIVIDUAL}$ -	12
	X_{SOCIAL} -	14	$X_{INDIVIDUAL}$ -	14
	Y_{SOCIAL} -	14	$Y_{INDIVIDUAL}$ -	14

Rewarded trials are denoted with (+) and unrewarded trials with (-). A and B refer to
pretraining trials, AX and BY to rewarded and CZ to unrewarded compound trials, and X
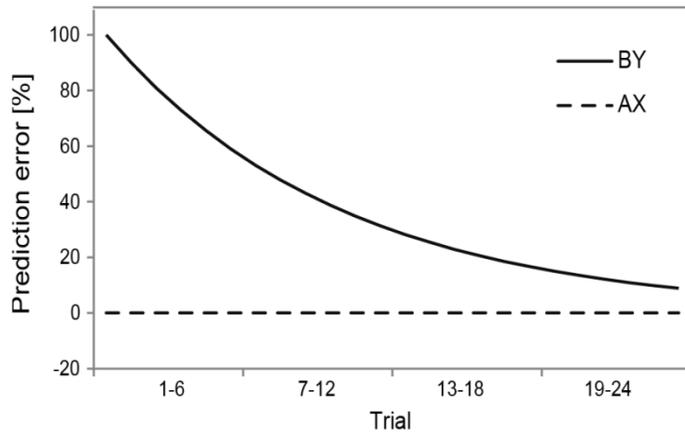and Y to unrewarded test trials.

Fig. S1 Theoretically expected prediction errors for BY and AX trials. At the beginning of the compound conditioning phase, BY trials elicit a positive prediction error, which gradually decreases over time as participants learn to associate BY with the reward. In contrast, the prediction error in AX trials should remain constantly close to zero as A trials have already been associated with reward during pretraining. Thus, in AX trials during compound conditioning, there should be no prediction error when rewards are shown. In our fMRI analysis, we captured this effect by contrasting parametric modulators corresponding to decreasing prediction error signals in BY versus AX trials. Thus, by contrasting parametric modulators, we asked which brain region shows a better fit with a decreasing prediction error signal in BY trials than it does in AX trials.
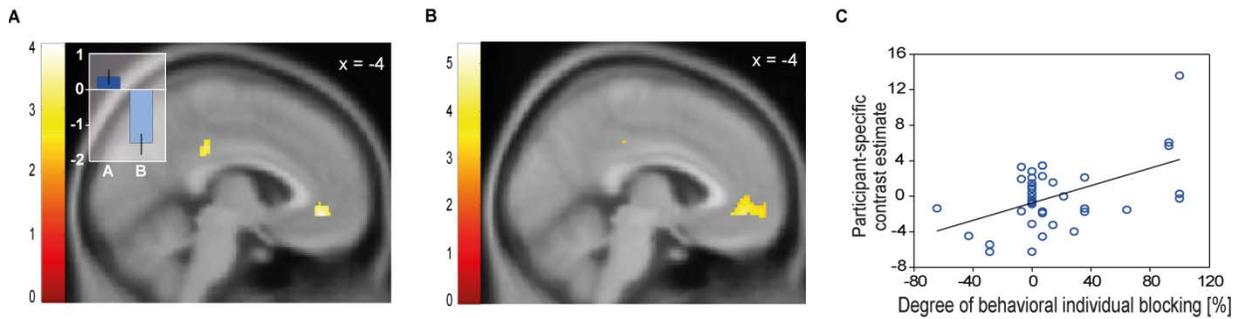
Fig. S2 Activity in vmPFC reflects expectation and blocking for individual rewards. *A*, vmPFC responses were higher to reward-predicting A stimuli as compared to neutral B stimuli (-4, 40, -6; $p < 0.05$, FWE small-volume corrected). Contrast estimates (inset) show vmPFC responses to A and B stimuli separately. Error bars indicate SEM. *B* and *C*, Differences in activation responses to Y stimuli as compared to X stimuli increased in the vmPFC (-6, 42, -4, $p < 0.05$, FWE small-volume corrected) with degree of behavioral blocking in the individual condition. Blocking was quantified as the difference in reward-expecting responses to non-blocked stimulus Y compared to blocked stimulus X.