

Helping or punishing strangers: Neural correlates of altruistic decisions as third-party and of its relation to empathic concern

Yang Hu, Sabrina Strang and Bernd Weber

Journal Name:	Frontiers in Behavioral Neuroscience
ISSN:	1662-5153
Article type:	Original Research Article
First received on:	16 Sep 2014
Revised on:	19 Jan 2015
Frontiers website link:	www.frontiersin.org

1 **Helping or punishing strangers: Neural correlates of altruistic** 2 **decisions as third-party and of its relation to empathic concern**

3

4 **Yang Hu^{1*†}, Sabrina Strang^{1,2†}, Bernd Weber^{1,3}**5 ¹Center for Economics and Neuroscience, University of Bonn, Bonn, Germany6 ²Department of Psychology, University of Lübeck, Lübeck, Germany7 ³Department of Epileptology, University Hospital Bonn, Bonn, Germany8 * **Correspondence:** Yang Hu, Center for Economics and Neuroscience, University of Bonn, Nachtigallenweg 86, Bonn,
9 53127, Germany.10 huyang@uni-bonn.de11 [†]These authors are co-first authors.12 **Keywords:** third-party help, third-party punishment, empathic concern, reward, striatum.

13

14 **Abstract**

15 Social norms are a cornerstone of human society. When social norms are violated (e.g. fairness)
16 people can either help the victim or punish the violator in order to restore justice. Recent research has
17 shown that empathic concern influences this decision to help or punish. Using functional magnetic
18 resonance imaging (fMRI) we investigated the neural underpinnings of third-party help and
19 punishment and the involvement of empathic concern. Participants saw a person violating a social
20 norm, i.e. proposing unfair offers in a dictator game, at the expense of another person. The
21 participants could then decide to either punish the violator or help the victim. Our results revealed
22 that both third-party helping as well as third-party punishing activated the bilateral striatum, a region
23 strongly related with reward processing, indicating that both altruistic decisions share a common
24 neuronal basis. In addition, also different networks were involved in the two processes compared
25 with control conditions; bilateral striatum and the right lateral prefrontal cortex (IPFC) during helping
26 and bilateral striatum as well as left IPFC and ventral medial prefrontal cortex (vmPFC) during
27 punishment. Further we found that individual differences in empathic concern influenced whether
28 people prefer to help or to punish. People with high empathic concern helped more frequently, were
29 faster in their decision and showed higher activation in frontoparietal regions during helping

30 compared with punishing. Our findings provide insights into the neuronal basis of human altruistic
31 behavior and social norm enforcement mechanism.

32 1. Introduction

33 Humans have an intriguingly complex social norm system, which is unique in the animal kingdom
34 and essential for the functioning of human society (Fehr & Rockenbach, 2004). Self-interests are
35 often in conflict with these social norms. When allocating resources our self-interests might lead us
36 to favor an unequal distribution at the expense of others, violating fairness or equal distribution
37 norms. When observing another person violating a social norm, e.g. treating another person unfair,
38 we have at least two options of how to react to this norm violation, namely to either punish the
39 offender, or to help (compensate) the victim. Punishing the offender is referred to as retributive
40 justice (Hogan & Emler, 1981) and helping the victim is referred to as compensatory justice (Darley
41 & Pittman, 2003). Usually people have to choose whom they want to focus on (i.e. the offender or
42 the victim) and then decide whether they want the offender to pay for what he or she did, or whether
43 they want to restore the harm done to the victim (Schroeder, Steel, Woodell, & Bembenek, 2003). It
44 was shown that people's first reaction to norm violations of high severity is to punish the offender.
45 However, people have a desire to help the victim after norm violations of low severity or when asked
46 to focus on the victim (Gromet & Darley, 2009). Furthermore, victims themselves attach importance
47 to being helped or compensated (Umbreit, 1998). Thus, both punishing the offender as well as
48 helping the victim are conceivable reactions to norm violations and might help to restore social
49 equity.

50 Helping a victim as well as punishing a norm violator as a third person (outside observer) can be
51 regarded as altruistic acts. Both cost people at least time and effort but provide no direct benefits.
52 Nevertheless people show altruistic helping (Leliveld, Van Dijk, & Van Beest, 2012) as well as
53 altruistic punishment (Fehr & Fischbacher, 2004; Fehr & Gächter, 2002). Both behaviors reduce
54 inequality between offender and victim. Recent neuroimaging studies suggest that altruistic behavior
55 is intrinsically rewarding as it was found to be correlated with activity in the striatum, an area known
56 to be involved in reward processing (Haber & Knutson, 2010). Specifically, the ventral striatum was
57 shown to be activated when people invest their own money to reduce their teammates' physical pain
58 (Hein, Silani, Preuschoff, Batson, & Singer, 2010) and when helping an African orphan (Genevsky,
59 Västfjäll, Slovic, & Knutson, 2013). Although the first party is not explicitly mentioned in these
60 studies, helping in this context can be regarded as a form of third party helping. Participants were not
61 involved in the unfair situation themselves (they were neither victims nor violators) but helped
62 another victim. De Quervain and colleagues found that the striatum was also involved in second party
63 punishment, namely when participants punished the untrustworthy opponent in a trust game
64 paradigm (de Quervain et al., 2004). In this case the participant was the victim of unfair behavior. So

65 far, there are only two studies on the neural correlates of third-party punishment (Buckholtz et al.,
66 2008; Strobel et al., 2011). In the study by Buckholtz and colleagues, participants were asked to rate
67 the appropriate punishment for crimes they were not involved in. However, decisions in this study
68 were not incentivized and therefore not costly for participants. In another study by Strobel and
69 colleagues, a modified paradigm of dictator game was adopted in which participants played the role
70 of either the recipient (i.e. second-party) or the observer (i.e. third-party) and they could punish the
71 dictator with their own money. They found that both second-/third-party punishment (vs. no
72 punishment) elicited stronger activation in ventral striatum. Thus, up to now neuroimaging studies
73 show that second-party punishment and third party help involve similar neuronal processes, namely
74 activity in reward areas. Reward might be an underlying mechanism for both third party help and
75 punishment decisions, thus both might involve activity in the striatum

76 Despite some similarities, behavioral studies suggest that third-party altruistic punishment and help
77 seem to be driven by different motives. On the one hand people feel sympathy/empathy with the
78 victim triggering a desire to restore the person (Gromet & Darley, 2009). On the other hand norm
79 violations induce strong negative affect which lead people to punish the offender (Egas & Riedl,
80 2008; Fehr & Gächter, 2002). One additional motive of punishment is deterrence; punishment has the
81 additional function to prevent offenders from future norm violations (Carlsmith, Darley, & Robinson,
82 2002). Taken together, behavioral studies suggest that third party help and punishment are differently
83 motivated and might therefore involve different processes. Intriguingly, people differ in their
84 responses when asked to choose between punishing the offender and helping the victim of a norm
85 violation. A recent behavioral study found that when witnessing an unfair case of monetary
86 allocation, people as third parties with low empathic concern preferred punishment, whereas those
87 high in empathic concern preferred helping (Leliveld et al., 2012). This indicates that empathic
88 concern plays an important role in influencing people's choice either to help or to punish. Empathic
89 concern is defined as an other-oriented altruistic motivation congruent with the perceived welfare of
90 another person; namely a feeling of concern for other people who are in need or suffer from an
91 unfortunate case (Batson et al., 1988; Coke, Batson, & McDavis, 1978). More crucially, previous
92 studies have shown empathic concern is a reliable indicator for helping behavior (Batson et al., 1988;
93 Coke, Batson, & McDavis, 1978). As a stable disposition variable, empathic concern was measured
94 by one subscale of the Interpersonal Reactivity Index (IRI_EC, German version; Davis, 1983). The
95 IRI_EC was also used in previous neuroimaging studies to investigate correlations between empathic
96 concern and empathic neural responses, however the results are inconclusive. One of the main
97 reasons is that different approaches were used in those studies to assess the neural correlates of
98 empathy, which makes it difficult to compare the results (Decety, 2011; Lamm, Nusbaum, Meltzoff,
99 & Decety, 2007; Singer et al., 2004). For example, Singer and colleagues (2004) adopted a cue-based
100 paradigm, in which participants' empathy was elicited by abstract visual information about their

101 partner's affective state. They found stronger positive relation between IRI_EC scores and empathic
102 neural activities in anterior cingulate cortex and left anterior insula. In the study by Lamm *et al.*
103 (2007), a picture-based paradigm was used, in which participants' empathy was elicited by viewing
104 other's body parts in painful situations (e.g. the painful needle injection on someone's hand).
105 However, no correlation was found between IRI_EC and empathic neural activities in those regions.

106 Although third-party help and punishment have been extensively investigated in behavioral studies,
107 the neuronal basis of third-party help and punishment has not been examined simultaneously in one
108 study using the same paradigm so far, allowing for a direct comparison. Furthermore, the association
109 between empathic concern and brain responses to third-party help or punishment is still unclear.
110 Adapting the paradigm of Leliveld *et al.* (2012) we investigated the neural correlates of third-party
111 help and punishment simultaneously in one study by using fMRI. Our aim was to examine the neural
112 processes underlying third-party help and punishment and their relation to individual differences in
113 empathic concern. Based on previous neuroimaging research we hypothesize that both third-party
114 help as well as punishing activates the striatum (de Quervain *et al.*, 2004; Genevsky *et al.*, 2013).
115 However, since behavioral studies showed that the motives to punish and help are different we
116 predict that help and punish elicit activity in separate brain regions connected to the striatum.
117 Furthermore, we assume that individual differences in empathic concern correlate with both the
118 frequency of help decisions and brain activity related to help (vs. punishment). Given that previous
119 studies do not report any consistent results about possible target regions for the connectivity analyses,
120 we refrain to make strong predictions but rather choose to present exploratory results.

121

122 **2. Materials and methods**

123 **2.1. Participants**

124 Thirty-six German participants (12 males; mean age = 22.72 ± 2.85) were tested in the fMRI
125 experiment. All participants reported no history of psychiatric or neurological disorders. They were
126 recruited via the Online Recruitment System for Economic Experiments (ORSEE). Written consent
127 was given by all participants according to the Declaration of Helsinki (BMJ 1991; 302: 1194) and the
128 study was approved by the ethics committee of the University of Bonn. Additional 84 participants
129 (30 males; mean age = 23.58 ± 6.13) were recruited for the behavioral experiment from the same
130 subject pool as used for the fMRI experiment.

131

132 **2.2. Stimuli and design**

133 The experiment consisted of two parts: a behavioral and an fMRI part. Participants of the behavioral
134 part were asked to play a Dictator Game. During ten rounds half of them played the role of the
135 “proposer” (i.e. first-party) and the other half the role of the “recipient” (i.e. second-party). We used a
136 perfect stranger matching to allocate participants for each round. The “proposer” received an
137 endowment of 100 monetary units (MUs; 1MU=20 Cents) per round and could decide how to
138 distribute these between him-/herself and the recipient (i.e. 0, 10, 20, 30, 40, 50). Participants were
139 informed that some of their decisions were forwarded to a third party (i.e. the fMRI participants). In
140 case of an unfair allocation the third party could decide to either help the recipient by transferring
141 MUs to increase his/her original MUs or to punish the proposer by investing own MUs to subtract
142 his/her original MUs. Participants were further asked to indicate their initials and were informed that
143 these were forwarded to the third parties. All participants of the behavioural experiment received a 4
144 € show-up fee at the end of the experiment. They were also informed that in addition all parties
145 would receive payoffs depending on one randomly chosen round of the experiment. Thus, if the third
146 party decided to either help the recipient or punish the dictator this decision was implemented
147 accordingly. The additional payoffs ($M = 10.05$ €, $SD = 7.26$ €) for participants of the behavioural
148 experiment were paid four weeks later. The behavioural part of experiment was conducted in Bonn
149 EconLab via Z-tree (Fischbacher, 2007).

150 In total, 420 decisions were made by the proposers, including 63 decisions of 50/50 offer, 43 offers
151 of 60/40 offer, 33 decisions of 70/30 offer, 57 decisions of 80/20 offer, 82 decisions of 90/10 offer
152 and 142 offers for 100/0. Given the goal of our study and the fMRI design, we focused on the unfair
153 offers (i.e. 60/40, 70/30, 80/20, 90/10, 100/0) and selected 160 offers to present those in the fMRI
154 study. Among them, 120 offers were presented in the decision condition and 40 in the control
155 condition. Each offer (i.e. 60/40, 70/30, 80/20, 90/10, 100/0) occurred 24 times in the decision
156 condition and 4 times in the control condition.

157

158 **2.3. fMRI procedure**

159 Participants were informed about the behavioral experiment and that they would see a set of
160 allocations made during this experiment. They were further told that they could influence the payoff
161 of either the first or second party by investing their own endowment. Importantly, both options were
162 costly for the participant, meaning that they had to invest one MU in order to either subtract three
163 from the proposer or to increase three to the recipient. Prior to the scanning session, participants
164 received an instruction which included a short comprehension test to further make sure that they
165 understood the task.

166 The scanning session consisted of two fMRI runs, which were separated by a self-paced break. In

167 each run, there were 80 trials; 60 decision trials (12 trials per offer) and 20 control trials (4 trials per
168 offer, half of them were in help/punish condition). In each trial, participants were endowed with 50
169 MU (1MU=20 Cents). In the decision condition, participants first saw the unfair monetary allocation
170 paired with the initials of the first and second party (**Figure 1A**). On the same screen they were asked
171 whether they wanted to increase the recipient's payoff or to decrease the proposer's payoff. Once
172 they made a choice, a cue appeared under the corresponding option (the decision phase). Independent
173 of their response time the decision phase was presented for 4s. The decision phase was followed by
174 an inter-stimulus fixation point (1-3s). On the next screen participants could decide how much they
175 want to increase or decrease the payoffs of the other players (the transfer phase; 4s), followed by an
176 inter-trial fixation point (3-7s). Participants could respond by pressing the button of response grips
177 with left/right index fingers in both phases of the task. In the control condition, the procedure was
178 identical except that in both phases decisions were made by the computer instead of the participants
179 lying in the scanner. The offers presented during these trials were still made by the participants of the
180 behavioral experiment though. Thus, participants in the scanner did not make any decisions
181 themselves, however, these trials were relevant for the payoffs of all parties (proposer, recipient and
182 fMRI participant). Participants therefore had an incentive to keep track of the control condition trials.
183 No button presses were asked of the participants in the control condition to limit the feeling of a
184 forced choice which might lead to conflict, anger or frustration. These trials were indicated by a
185 white frame (**Figure 1B**). The display of the task and response collection was performed with
186 Presentation 14.9 (Neurobehavioral System, Albana, Canada). Participants saw the experiment via
187 video goggles (Nordic NeuroLab, Bergen, Norway) and their responses were recorded by response
188 grips (Nordic NeuroLab, Bergen, Norway).

189

(Insert **Figure 1** about here)

190 It is important to highlight the following details of the paradigm and the procedure. First, the
191 words "help" and "punish" were not used in the instructions ("increase" and "subtract" were adopted
192 instead) to avoid demand characteristics. Second, consistent with previous literature (Fehr &
193 Fischbacher, 2004; Leliveld et al., 2012), the cost ratio was set to 1:3, which means that 1 MU
194 transferred from participants could either subtract 3 MU from the first party or increase 3 MU to the
195 second party. Third, in the transfer phase participants could decide to invest 0 MU. Thus every
196 decision to invest MU to either increase or decrease MUs of the others can be regarded as their
197 voluntary decision. Fourth, the position of two options (i.e. "increase" and "subtract") in the decision
198 phase were counterbalanced across trials. The default position of the amount participants could invest
199 in the transfer phase was randomly determined from 0 to 50. Finally, the first party could not lose
200 money (i.e. the minimum payoff was 0).

201 After scanning, participants were asked to fill in the Interpersonal Reactivity Index (IRI) scale,

202 used for measuring trait empathy and to make judgments about the fairness of the six different offers
203 (i.e. the offer 50/50 was also included) on a 8-point Likert scale (1=very fair, 8=very unfair). Finally,
204 participants received a 10 € show-up fee and one randomly selected trial was paid to all three parties
205 accordingly ($M = 7.0$ €, $SD = 2.5$ €).

206

207 **2.4. Data collection and analyses**

208 The imaging data was collected via the 3-Tesla Siemens Trio platform at the Imaging Center of Life
209 & Brain, University Hospital Bonn. For functional images, 37 axial slices ($FOV = 192 \times 192$ mm²,
210 matrix = 96×96 , in-plane resolution = 2×2 mm², thickness = 3 mm) covering the whole brain were
211 obtained using a T2*-weighted echo planar imaging (EPI) sequences with blood-oxygenation-level
212 dependent (BOLD) contrast (TR = 2500 ms, TE = 30 ms, flip angle = 90 °). A high-resolution
213 structural image for each participant was acquired using 3D MRI sequences for anatomical co-
214 registration and normalization (TR = 1660 ms, TE = 2.75 ms, flip angle = 9 °, matrix = 320×320 ,
215 $FOV = 256 \times 256$ mm², slice thickness=0.8 mm).

216 Eleven participants were excluded due to the following reason: 10 of them had insufficient number
217 of trials (less than 5 trials) for one or both decision regressors (help decision: $n=1$; punish decision:
218 $n=7$; both decisions: $n = 2$) and one participant terminated the experiment because he or she felt
219 uncomfortable in the scanner. For the remaining 25 participants, SPM8 was used for the fMRI data
220 analysis (Wellcome Trust Department of Cognitive Neurology, London, UK). For each run of each
221 participant, the first three volumes were discarded to allow the stabilization of BOLD signal. The
222 following preprocessing steps were applied: EPI images were first realigned to the first volume to
223 correct for head motions (< 2.5 mm) and corrected for slice timing. Then, the anatomical image was
224 co-registered to the mean EPI image, and segmented, generating parameters for normalization to
225 MNI space. Using these parameters, all EPI data were projected onto MNI space with a $2 \times 2 \times 2$ mm³
226 resolution and smoothed using an 8-mm FWHM (full width half maximum) isotropic Gaussian
227 kernel. High-pass temporal filtering with a cut-off of 128 s was performed to remove low-frequency
228 drifts.

229 For the individual-level analyses, a general linear model (GLM) focusing on the decision-phase
230 with five onset regressors (i.e., “help”, “punish”, “help_control”, “punish_control”, “other”)
231 convolved with the canonical hemodynamic response function (HRF) was applied. The “other”
232 regressor included the following onsets: onsets of transfer phase and onsets of no response as well as
233 trials in which participants transferred 0 MU in decision phase. For runs in which either “help” or
234 “punish” condition was less than 5 trials, onsets of that condition in decision phase were also
235 categorized into “other” condition. The six estimated head movement parameters were included in

236 the design matrix to account for the residual effects of head motion. For the group-level analyses, a
237 one-sample t-test as well as a flexible factorial model was performed to test the difference and the
238 conjunction of the activation elicited by “help” and “punish” option. Parameter estimates (contrast
239 values) and percent signal change of the peak voxel was extracted via MarsBar
240 (<http://marsbar.sourceforge.net>).

241

242 **2.4.1. Correlation analysis**

243 To investigate how trait empathy correlates with third-party decisions at the neural level, a
244 correlation analysis was applied to compute the relationship between the individual neural contrast of
245 “help” vs. “punish” and individual scores of empathic concern subscale of IRI (IRI_EC).

246

247 **2.4.2. Psycho-physiological interaction (PPI) analysis**

248 In order to test whether different networks are involved during helping and punishing respectively,
249 we performed a PPI analysis (Friston *et al.*, 1997; Gitelman, Penny, Ashburner, & Friston, 2003).
250 Specifically, the source masks were defined as two 8-mm spheres centered at the peak voxel of the
251 group-level conjunction results of the two contrasts “help” vs. “help_control” and “punish” vs.
252 “punish_control” within bilateral striatum based on AAL templates with the *wfu_pickatlas* tool. The
253 seed volume of interest (VOI) for each individual was then defined as a sphere with a 6-mm-radius
254 centered at the peak voxel from the contrast of either “help” vs. “help_control” or “punish” vs.
255 “punish_control” within these source masks. The time series of each VOI was extracted and then
256 deconvolved, multiplied with the psychological variable (“help” > “help_control” or “punish” >
257 “punish_control”) and reconvolved with a hemodynamic response function to set up the PPI
258 regressor, which followed the procedure by Gitelman *et al.* (2003). These three regressors (i.e. the
259 PPI regressor, the VOI time-series, the psychological variable) were convolved with the canonical
260 HRF and then entered into the regression model along with six head motion parameters. The
261 individual parameter estimates image for the PPI regressor was subsequently subjected to one-sample
262 t-tests. Finally, a group analysis was performed to identify the brain regions displaying increased
263 functional connectivity with the seed VOI during either help or punishment decisions. Besides, two
264 paired-samples t-tests were performed to further test the different connectivity patterns between help
265 and punishment decisions with either left or right striatum.

266 For all whole-brain based analyses mentioned above, the threshold of $p < 0.001$ uncorrected at
267 peak voxel level with the extent threshold at $k = 50$ was adopted.

268

269 **3. Results**270 **3.1. Behavioral results**

271 Data from 25 participants were used for behavioral analyses. A paired-samples T-test was performed
272 between help and punishment decisions in the decision condition on the behavioral factors *decision*
273 *rate* (i.e. the ratio of help/punish decision compared in relation to all respective trials), *response time*
274 (ms) and *transfer amount* (MU). The *transfer amount* was significantly different between help and
275 punishment trials. Participants transferred more MUs when they punished the first party (M = 16.15,
276 SD = 6.86) than when they helped the second party (M = 11.07, SD = 5.07) [95% C.I. of the
277 difference: -8.28 to -1.89; $t_{(24)}=3.266$, $p = 0.003$, Cohen'd = -0.664]. No significant differences were
278 detected in the rate of decisions to help or punish (help: M = 49.30%, SD = 27.28%; punish: M =
279 42.40%, SD = 27.90%) [95% C.I. of the difference: -15.62% to 29.42%; $t_{(24)}=0.632$, $p = 0.533$,
280 Cohen'd = 0.126] and *response times* (help: M = 1583.15, SD = 431.63; punish: M = 1611.45, SD =
281 402.22) [95% C.I. of the difference: -207.55 to 150.93; $t_{(24)}=-0.326$, $p = 0.747$, Cohen'd = -0.065]
282 between help and punishment trials.

283 To test whether individual differences in trait empathy correlate with the decisions to help or
284 punish, a Pearson correlation was conducted between empathic concern subscale scores of IRI (i.e.
285 IRI_EC) and *decision rate* in help and punishment decisions respectively. A significant positive
286 relationship was found between IRI_EC scores and *help rate* [95% C.I.: 0.06 to 0.71; $r = 0.441$, $p =$
287 0.027 , Fisher's $Z_r = 0.474$], whereas a negative relationship was detected between IRI_EC scores and
288 *punishment rate* [95% C.I.: -0.72 to -0.08; $r = -0.461$, $p = 0.02$, Fisher's $Z_r = -0.497$; **Figure 2A**]. To
289 further investigate whether empathic concern has an influence on decision speed in both help and
290 punishment trials, we correlated IRI_EC and the difference in reaction times between help and
291 punishment trials (i.e. RT_help-punish), finding a negative relationship [95% C.I.: -0.69 to -0.01; $r =$
292 -0.406 , $p = 0.044$, Fisher's $Z_r = -0.431$; **Figure 2B**].

293

(Insert **Figure 2** about here)

294 A one-way repeated measure ANOVA on the perceived unfairness rating of the offers showed a
295 main effect of inequity level [95% C.I.: 5.09 to 5.51; $F_{(5, 120)} = 225.967$, $p < 0.001$, partial $\eta^2 = 0.904$].
296 Post-hoc analyses revealed that ratings increased with the level of inequity of the offers (50/50: M =
297 1.48, SD = 1.12; 60/40: M = 3.52, SD = 1.30; 70/30: M = 5.24, SD = 0.93; 80/20: M = 6.24, SD =
298 0.93; 90/10: M = 7.32, SD = 0.48; 100/0: M = 8.00, SD = 0.00; $ps < 0.05$, *Bonferroni* corrected).

299

300 3.2. Imaging findings

301 3.2.1. Neural correlates of third-party help and punishment

302 Both contrast *help* vs. *help_control* and *punish* vs. *punish_control* showed significant activation in
303 several regions, including bilateral striatum, supplementary motor area/mid-cingulate cortex (BA 4/6),
304 inferior/superior parietal lobule (BA 39/40) as well as visual areas (BA 17/18/19) (**Table S1** and
305 **Figure 3**). The conjunction analyses further confirmed that the bilateral striatum along with other
306 areas mentioned above were activated by both contrasts, indicating that help- and punish-related
307 cognitive processes shared some common neural bases (**Table S1** and **Figure 3**). Activity in the
308 bilateral striatum remained significant when controlling for motor responses due to button pressing
309 (**Table S2** and **Figure S1**). To test the differential neural correlates between these two processes, help
310 and punishment decisions were directly contrasted, which yielded no significant difference in both
311 directions. These results remained unchanged when controlling for fairness levels and transfer
312 amounts (**Table S3** and **S4**).

313 (Insert **Figure 3** about here)

315 3.2.2. Relationship between empathic concern and brain activation during third-party 316 decisions

317 To determine regions in which a change of the BOLD signal to third-party decisions varied with
318 individual difference in trait empathy, a correlation analysis was performed between the contrast of
319 help vs. punishment and IRI_EC scores. Stronger positive correlations were detected in fronto-
320 parietal regions including left lateral prefrontal cortex (IPFC, BA 9) and left angular gyrus/inferior
321 parietal lobule (IPL/AG, BA 7/40; **Table S5** and **Figure 4**). No negative correlations were found
322 under the same threshold.

323 (Insert **Figure 4** about here)

325 3.2.3. Functional connectivity pattern of third-party decisions

326 In order to investigate whether different networks are involved in third-party help and punishment a
327 PPI analysis was conducted. Based on our hypotheses and the results of the conjunction analyses the
328 striatum was used as the seed region (i.e. left and right striatum). PPI analyses were conducted during
329 help and punishment decisions, respectively (both compared with their respective control conditions).
330 Right IPFC (BA 45/46) showed increased functional connectivity with bilateral striatum during help

331 decisions (**Table S6** and **Figure 5**), whereas left IPFC (BA 44/45) showed enhanced functional
332 connectivity with both seed regions during punishment decisions (**Table S6** and **Figure 6**).
333 Furthermore ventral medial prefrontal cortex (vmPFC; BA 10/11/32) was observed to show increased
334 connectivity only with right striatum when participants chose to punish (**Table S6** and **Figure 6**). No
335 significant difference in functional connectivity was found in a direct comparison of help and
336 punishment decisions with either left or right striatum.

337 (Insert **Figure 5** about here)

338 (Insert **Figure 6** about here)

339

340 **4. Discussion**

341 Our results reveal that both third-party help and third-party punishment share a common neuronal
342 basis, but that specific networks are additionally involved in the two processes. The bilateral striatum
343 was activated by both helping and punishing; functional connectivity between the bilateral striatum
344 and the right lateral prefrontal cortex (IPFC) was increased during help and with left IPFC and
345 ventromedial prefrontal cortex (vmPFC) during punishment. Individual differences in empathic
346 concern correlated with people's preference to help or to punish. People with high empathic concern
347 helped more frequently, were faster in their decision and showed higher activation in fronto-parietal
348 regions during decisions to help.

349 The conjunction analysis indicated that third-party help and third-party punishment both share
350 some common neural bases. In line with previous findings the striatum showed increased activation
351 during altruistic help (Genevsky et al., 2013; Harbaugh, Mayr, & Burghart, 2007; Hein et al., 2010)
352 as well as during altruistic punishment (de Quervain et al., 2004). Helping friends or even strangers
353 and punishing norm violators has been associated with activity in the striatum. However, so far
354 striatal activation was only observed in third party helping and second-party punishment paradigms,
355 for example, while an investor chose to punish an untrustworthy trustee (de Quervain et al., 2004).
356 This is to our knowledge the first study investigating third party helping and punishing in the same
357 study and showing that both are associated with striatal activation. The striatum is part of the human
358 reward system, known to be activated by recognizing and evaluating rewards and learning from them
359 (Bhanji & Delgado, 2014). Our results are in line with literature on charitable donation and second
360 party punishment suggesting that both helping an unknown person and punishing an offender is
361 intrinsically rewarding (Fehr & Camerer, 2007; Harbaugh et al., 2007). Here we show that punishing
362 an offender as a third person seems to rewarding as well.

363 However, an alternative interpretation of this result cannot be ruled out completely. Participants
364 were not required to respond during the computer (control) trials in order to avoid additional
365 cognitive (e.g. conflict) or affective (e.g. anger, frustration) processes. Unfortunately this paradigm
366 introduced a potential motor confound for the contrasts between help or punish decisions (button
367 presses) and their corresponding control trials (no button presses). Besides its role in reward
368 processing or representation of affective value, the striatum is also frequently associated with motor-
369 related functions (Witt, Laird, & Meyerand, 2008; Filevich, Kühn, & Haggard, 2012; Guitart-Masip
370 et al., 2012; Guitart-Masip et al., 2014). In a recent study on the role of the striatum in decision
371 making, Guitart-Masip and colleagues independently manipulated both the variables of action (i.e.
372 “go” or “no go”) and valence (i.e. “to win” or “to avoid losing”) in an instrumental learning paradigm.
373 They found that activity in the striatum reflected primarily the action requirements, independent of
374 the valence of decision (Guitart-Masip et al., 2012). This result suggests an involvement of the
375 striatum in motivated action during decision making. In order to control for this, we performed an
376 additional analysis in which the onset of the button presses were added in to the GLM as an
377 independent regressor. This analysis showed that the bilateral striatum was still strongly activated
378 during both third-party help and punishment even after controlling for the effect of button pressing
379 (see **Table S2** and **Figure S1**), indicating that activity in the striatum detected in the contrasts of
380 third-party altruistic decisions and control trials is not likely driven by pure motor effects only.
381 Rather it more likely reflects processes related to decision making, like rewarding processes as
382 suggested by previous findings on altruistic decisions (e.g. charity donation, second-/third-party
383 punishment; de Quervain et al., 2004; Harbaugh et al., 2007; Hein et al., 2010; Strobel et al., 2011;
384 Genevsky et al., 2013) and on reward processing (Haber & Knutson, 2010). However, since the onset
385 of button pressing is not a random event as it is collinear to the onset of decision trials, the analysis
386 unfortunately might not completely tease apart the effect of button pressing and that of decision
387 processes. Since we cannot perfectly disentangle brain activity due to decision processes and due to
388 motor processes (button press), the joint activation in striatum during third-party help and
389 punishment decisions should be cautiously interpreted as reward-relevant processing.

390 Furthermore, our functional connectivity results suggest that besides the common neural basis,
391 different networks are involved in third-party help and third-party punishment. Increased functional
392 connectivity was found between the bilateral striatum and right IPFC during help decisions whereas
393 left IPFC and the bilateral striatum showed increased functional connectivity during punishment
394 decisions. Furthermore, vmPFC showed increased connectivity with right striatum when participants
395 chose to punish. Generally, our PPI findings are consistent with the anatomical connectivity of the
396 striatum, which was found to be connected with both lateral and ventral/medial parts of the prefrontal
397 cortex (Haber & Knutson, 2010). Specifically, IPFC is known to be engaged in cognitive/executive
398 control and goal-directed decisions (Miller & Cohen, 2001; Tanji & Hoshi, 2008). In the social-

399 economic domain, especially the right IPFC was shown to be involved in the control of selfish
400 impulses (Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; Ruff, Ugazio, & Fehr, 2013; Strang
401 et al., 2014). For example, disrupting the right IPFC via lower-frequency repetitive TMS caused
402 people to make riskier decisions (Knoch & Fehr, 2007) and to exhibit more norm violating
403 behaviours (Strang et al., 2014). A recent TMS study found that people show more impulsive
404 behavior in an inter-temporal choice task while the left IPFC was inhibited (Figner et al., 2010).
405 Intriguingly, left IPFC showed stronger activity when participants chose to costly punish dictators as
406 a third-party compared to as a second-party, indicating that cognitive-control processes, instead of
407 revenge-driven motives, are involved in third-party punishment (Strobel et al., 2011). Consistently, a
408 behavioral study showed that punishment of free-riders by cooperators is linked to self-control
409 abilities (Espín, Brañas-Garza, Herrmann, & Gamella, 2012). Moreover, some studies showed an
410 increased functional connectivity between IPFC and striatum while people controlled reward-related
411 responses to food cues (Hare, Malmaud, & Rangel, 2011) or monetary reward (Delgado, Gillis, &
412 Phelps, 2008). Since help and punishment are costly in our paradigm, both require control of selfish
413 impulses in order to engage in one of the two behaviors. Hence, it is possible that in our paradigm
414 increased connectivity between IPFC and bilateral striatum during help and punishment decisions is
415 due to these control processes. Another region found to have increased connectivity with the striatum
416 is the vmPFC, which has been shown to be involved in a variety of cognitive and affective processes
417 including integrating emotional information (Naqvi, Shiv, & Bechara, 2006) and subjective valuation
418 during decision making (Ruff & Fehr, 2014). Interestingly, increased activity in the vmPFC was also
419 found when choosing to costly punish an untrustworthy trustee (de Quervain et al., 2004). Our
420 findings seem to support the view of a potentially stronger involvement of the vmPFC in third-party
421 punishment rather than help.

422 Nevertheless, two issues limit our interpretation on the PPI results. Firstly, it is important to
423 mention that there was no significant difference between the functional connectivity during help and
424 punishment decisions when directly contrasting the connectivity results in both decisions, which
425 weakens our inference about differential neural networks involved in each decision. This might be
426 due to insufficient sample size or inadequate numbers of trials in each condition. Secondly, as both
427 PPI analyses are based on the corresponding control trials (help vs. help_control or punish vs.
428 punish_control), the results are also influenced by the motor confound mentioned above. Thus, the
429 connectivity pattern might also reflect a motor effect during both decisions compared with the pure
430 observation in the control trials. Since the PPI analyses are rather explorative, further research is
431 needed to shed more light on the network involved in third party help and punishment.

432 Moreover, our results demonstrate that individual differences in empathic concern influence our
433 decision to help or to punish on a behavioral as well as on a neural level. People with high levels of

434 empathic concern chose to help more frequently, were faster in their decision to help and showed
435 higher activation in frontoparietal regions (i.e. left IPFC and left IPL/AG) during this decision. The
436 behavioral findings are in line with previous research (Leliveld et al., 2012), in which the authors
437 also reported that people with high empathic concern prefer to help instead of punishing, whereas
438 people with low empathic concern prefer to punish instead of helping. In addition our results
439 demonstrate that high empathic people are also faster in deciding to help compared to deciding to
440 punish, whereas people with low empathic concern show the reversed pattern; they are faster in
441 deciding to punish instead of helping. Faster reaction times are often interpreted as a sign of less
442 conflict between the options someone has to choose from and less cognitive processing (Rand,
443 Greene, & Nowak, 2012). According to this literature the results suggest that for high empathic
444 people deciding to help needs less cognitive processing. For them the decision to either help or to
445 punish does not involve a conflict, help is the default option for them. Low empathic people also do
446 not encounter a conflict when deciding between help and punishment, their default option is to
447 punish. Whether someone helps a victim or punishes the offender hence depends on how much
448 empathic concern someone has. Both regions correlating with empathic concern, IPFC and IPL/AG,
449 are considered as the core components of the frontoparietal network (FPN), which play an important
450 role in top-down cognitive control and attention (Corbetta & Shulman, 2002; Dosenbach, Fair, Cohen,
451 Schlaggar, & Petersen, 2008). Gromet and Darley (2009) argue that punishment might be the default
452 choice after observing injustice until people are asked to focus on the victim. Without explicit
453 requirements to focus on the victim, such an attention shift might be influenced by individual's
454 personality trait, in this case empathic concern. Our results hint towards such an empathy-based
455 attention shift. However, this interpretation is inconsistent with the reaction time findings, which
456 suggests that help is the default for people with high empathic concern. Thus the role of FPN in
457 mediating the relationship between empathic concerns and the two altruistic decisions still needs
458 further investigation. Future studies might shed more light on this question by adopting other
459 techniques (such as eye-tracking) to investigate the difference in fine-grained information search
460 patterns between high and low empathic people during deciding to either help or punish.

461 There are several limitations of this study. One constraint is the difference in motor demands
462 between the decision and control trials as mentioned above. Future studies should try to find a clearer
463 way to disentangle activity due to the decision process and motor responses. Another limitation is the
464 high number of excluded participants. We were only able to use data from 25 out of 36 participants,
465 because ten participants did not show enough variability in their behavior to define all necessary
466 regressors. Since trials were sorted into different conditions according to participant's behavior in the
467 corresponding trial, sufficient numbers of trials (>25) for one condition in order to calculate a
468 contrast cannot be guaranteed. Although 25 participants is still a widely accepted sample size in the
469 field of cognitive neuroimaging, statistical power might explain the non-significant difference

470 especially for the PPI results. Since people who exhibit either very high or very low empathic
471 concerns have a preference for either helping or punishing, respectively, they show less variability in
472 their decisions on the individual level. One possibility to minimize dropout rates is to increase the
473 variability in decision behavior by only inviting participants with empathic concern score in the
474 medium range and thereby increasing statistical power. Additionally a pre-screening could be used to
475 exclude participants who are very selfish and are not willing to help or punish at all.

476 Taken together, by using a modified third-party decision paradigm with fMRI, our study provides
477 first evidence for the neural basis of third-party help and punishment decisions. Both altruistic
478 decisions activated bilateral striatum, indicating that intrinsic reward processes are involved in both
479 third-party help and punishment decisions. Differential functional connectivity networks during
480 third-party help and punishment suggest different cognitive processes underlying both decisions.
481 Moreover, the present study replicated previous behavioral findings on the role of empathic concern
482 in mediating people's decisions to either help or punish. Further its underlying neural correlates in
483 frontoparietal regions were detected. Despite some limitations, these findings may provide insights
484 for better understanding the mechanism underlying altruism and social norm enforcement.

485

486 5. Acknowledgement

487 We would like to thank Sandra Mastani for her assistance in data collection. This work was
488 supported by the State Scholarship Fund of the China Scholarship Council (CSC; No. 201306140034)
489 to Y.H. and a Heisenberg Grant of the German Research Council (DFG; We 4427/3-1) to B.W.

490

491 6. References¹

492 Batson, C. D., Quin, K. O., Fultz, J., & Vanderplas, M. (1983). Influence of Self-Reported Distress
493 and Empathy on Egoistic Versus Altruistic Motivation to Help. *Journal of Personality and Social*
494 *Psychology*, 45(3), 706–718.

495 Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., ... Wicklund,
496 R. (1988). Five Studies Testing Two New Egoistic Alternatives to the Empathy-Altruism Hypothesis,
497 (1), 52–77.

¹ Provide the doi when available, and ALL complete author names.

- 498 Bhanji, J. P., & Delgado, M. R. (2014). The Social Brain and Reward: Social Information Processing
499 in the Human Striatum. Wiley Interdisciplinary Reviews. *Cognitive Science*, 5(1), 61–73.
- 500 Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the
501 motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, 34(4),
502 537-541.
- 503 Buckholz, J. W., Asplund, C. L., Dux, P. E., Zald, D. H., Gore, J. C., Jones, O. D., & Marois, R.
504 (2008). The neural correlates of third-party punishment. *Neuron*, 60(5), 930–940.
- 505 Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just
506 deserts as motives for punishment. *Journal of Personality and Social Psychology*, 83(2), 284–299.
- 507 Coke, J., Batson, C., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model.
508 *Journal of Personality and Social Psychology*, 36(7), 752–766.
- 509 Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the
510 brain. *Nature Reviews. Neuroscience*, 3(3), 201–215.
- 511 Darley, J. M., & Pittman, T. S. (2003). The Psychology of Compensatory and Retributive Justice.
512 *Personality and Social Psychology Review*, 7(4), 324–336.
- 513 Davis, M. H. (1983). Measuring Individual Differences in Empathy: Evidence for a
514 Mutlidimensional Approach. *Journal of Personality and Social Psychology*, 44(1), 113–126.
- 515 De Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., &
516 Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305(5688), 1254–1258.
- 517 Decety, J. (2011). Dissecting the Neural Mechanisms Mediating Empathy. *Emotion Review*, 3(1),
518 92–108.
- 519 Delgado, M. R., Gillis, M. M., & Phelps, E. (2008). Regulating the expectation of reward via
520 cognitive strategies. *Nature Neuroscience*, 11(8), 880–881.
- 521 DellaVigna, S., List, J., & Malmendier, U. (2012). Testing for altruism and social pressure in
522 charitable giving. *The Quarterly Journal of Economics*, 127(1), 1–56.
- 523 Desmond, J. E., Gabrieli, J. D., Wagner, A. D., Ginier, B. L., & Glover, G. H. (1997). Lobular
524 patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by

- 525 functional MRI. *The Journal of Neuroscience*, 17(24), 9675-9685.
- 526 Dosenbach, N. U. F., Fair, D., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-
527 networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3), 99–105.
- 528 Dunn, E. W., Aknin, L. B., & Norton, M. I. (2008). Spending money on others promotes happiness.
529 *Science*, 319(5870), 1687–1688.
- 530 Egas, M., & Riedl, A. (2008). The economics of altruistic punishment and the maintenance of
531 cooperation. *Proceedings of Biological Sciences*, 275(1637), 871–878.
- 532 Eisenberg, N., & Miller, P. (1987). The Relation of Empathy to Prosocial and Related Behaviors.
533 *Psychological Bulletin*, 101(1), 91–119.
- 534 Esp ñ, A., Bra ñas-Garza, P., Herrmann, B., & Gamella, J. (2012). Patient and impatient punishers of
535 free-riders. *Proceedings of Biological Sciences*, 279(1749), 4923–4928.
- 536 Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human*
537 *Behavior*, 25(2), 63–87.
- 538 Fehr, E., & G ächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.
- 539 Fehr, E., & Rockenbach, B. (2004). Human altruism: economic, neural, and evolutionary
540 perspectives. *Current Opinion in Neurobiology*, 14(6), 784–790.
- 541 Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., & Weber, E. U. (2010).
542 Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience*, 13(5), 538–
543 539.
- 544 Filevich, E., Kühn, S., & Haggard, P. (2012). Intentional inhibition in human action: the power of
545 ‘no’. *Neuroscience & Biobehavioral Reviews*, 36(4), 1107-1118.
- 546 Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental*
547 *Economics*, 10(2), 171–178.
- 548 Friston, K., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997). Psychophysiological and
549 modulatory interactions in neuroimaging. *Neuroimage*, 6(3), 218-229.
- 550 Genevsky, A., V ästf j äll, D., Slovic, P., & Knutson, B. (2013). Neural underpinnings of the
551 identifiable victim effect: affect shifts preferences for giving. *The Journal of Neuroscience*, 33(43),

- 552 17188–14196.
- 553 Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling regional and
554 psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution.
555 *Neuroimage*, 19(1), 200-207.
- 556 Gromet, D. M., & Darley, J. M. (2009). Punishment and beyond: Achieving Justice through the
557 Satisfaction of Multiple Goals. *Law & Society Review*, 43(1), 1–38.
- 558 Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision
559 making. *Trends in Cognitive Sciences*, 18(4), 194-202.
- 560 Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and
561 no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage*, 62(1),
562 154-166.
- 563 Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging.
564 *Neuropsychopharmacology*, 35(1), 4–26.
- 565 Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary
566 giving reveal motives for charitable donations. *Science*, 316(5831), 1622–1625.
- 567 Hare, T. a, Malmaud, J., & Rangel, A. (2011). Focusing attention on the health aspects of foods
568 changes value signals in vmPFC and improves dietary choice. *The Journal of Neuroscience*, 31(30),
569 11077–11087.
- 570 Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup
571 and outgroup members' suffering predict individual differences in costly helping. *Neuron*, 68(1),
572 149–160.
- 573 Hogan, R., & Emler, N. (1981). *The Justice Motive in Social Behavior*. (M. J. Lerner & S. C. Lerner,
574 Eds.) (pp. 125–143). Boston, MA: Springer US.
- 575 Knoch, D., & Fehr, E. (2007). Resisting the power of temptations: the right prefrontal cortex and
576 self-control. *Annals of the New York Academy of Sciences*, 1104, 123–134.
- 577 Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal
578 fairness by disrupting the right prefrontal cortex. *Science*, 314(5800), 829–832.

- 579 Lamm, C., Nusbaum, H. C., Meltzoff, A. N., & Decety, J. (2007). What are you feeling? Using
580 functional magnetic resonance imaging to assess the modulation of sensory and affective responses
581 during empathy for pain. *PloS One*, 2(12), e1292.
- 582 Leliveld, M. C., Van Dijk, E., & Van Beest, I. (2012). Punishing and compensating others at your
583 own expense: The role of empathic concern on reactions to distributive injustice. *European Journal*
584 *of Social Psychology*, 140, 135–140.
- 585 Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual*
586 *Review of Neuroscience*, 24, 167–202.
- 587 Naqvi, N., Shiv, B., & Bechara, A. (2006). The Role of Emotion in Decision Making: A Cognitive
588 Neuroscience Perspective. *Current Directions in Psychological Science*, 15(5), 260–264.
- 589 Rand, D. G., Greene, J. D., & Nowak, M. a. (2012). Spontaneous giving and calculated greed. *Nature*,
590 489(7416), 427–430.
- 591 Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making.
592 Nature Reviews. *Neuroscience*, 15(8), 549–562.
- 593 Ruff, C. C., Ugazio, G., & Fehr, E. (2013). Changing social norm compliance with noninvasive brain
594 stimulation. *Science*, 342(6157), 482–484.
- 595 Schroeder, D. A., Steel, J. E., Woodell, A. J., & Bembenek, A. F. (2003). Justice Within Social
596 Dilemmas. *Personality and Social Psychology Review*, 7(4), 374–387.
- 597 Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for
598 pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157–1162.
- 599 Strang, S., Gross, J., Schuhmann, T., Riedl, A., Weber, B., & Sack, A. T. (2014). Be Nice if You
600 Have to - The Neurobiological Roots of Strategic Fairness. *Social Cognitive and Affective*
601 *Neuroscience*. doi:10.1093/scan/nsu114
- 602 Strobel, A., Zimmermann, J., Schmitz, A., Reuter, M., Lis, S., Windmann, S., & Kirsch, P. (2011).
603 Beyond revenge: neural and genetic bases of altruistic punishment. *NeuroImage*, 54(1), 671–680.
- 604 Tanji, J., & Hoshi, E. (2008). Role of the lateral prefrontal cortex in executive behavioral control.
605 *Physiological Reviews*, 88(1), 37–57.

- 606 Umbreit, M. (1998). Restorative Justice Through Victim-Offender Mediation: A Multi-Site
607 Assessment. *Western Criminology Review*, 1. Retrieved from
608 <http://wcr.sonoma.edu/v1n1/umbreit.html>
- 609 Witt, S. T., Laird, A. R., & Meyerand, M. E. (2008). Functional neuroimaging correlates of finger-
610 tapping task variations: an ALE meta-analysis. *NeuroImage*, 42(1), 343-356.

611 7. Figure legends

612 **Figure 1 Illustration of trial procedures (A) in the decision condition and (B) in the control**
613 **condition.** ISI = inter-stimulus interval; ITI = inter-trial interval.

614 **Figure 2 Correlation (A) between IRI_EC scores (X-axis) and average help/punish rate (Y-axis)**
615 **and (B) between IRI_EC scores (X-axis) and the difference in RT between help and punish (Y-**
616 **axis).** IRI_EC = empathic concern subscale of interpersonal reactivity index scale; RT = response
617 time.

618 **Figure 3 Separate and conjunction mapping of regions involved in third-party help and**
619 **punishment (A) and timecourse of percent signal change in the local peak voxel of left striatum**
620 **(B) and right striatum (C) in four conditions** (i.e. help, help_control, punish, punish_control).
621 Error bars: SEM.

622 **Figure 4 Correlation between the contrast of help vs. punish and IRI_EC scores (A) and plots**
623 **of the positive correlation between IRI_EC scores and contrast values in local peak voxel of left**
624 **IPFC (B) and that of left IPL/AG (C).** IRI_EC = empathic concern subscale of interpersonal
625 reactivity index scale; IPFC = lateral prefrontal cortex; IPL = inferior parietal lobule; AG = angular
626 gyrus.

627 **Figure 5 Regions showing increased functional connectivity with bilateral striatum during**
628 **third-party help decisions (compared with control conditions; A) and plots of parameter**
629 **estimates of PPI in the local peak voxel of right IPFC with left (B) /right (C) striatum in four**
630 **conditions** (i.e. help, help_control, punish, punish_control). Abbreviations: PPI = psycho-
631 physiological interaction; IPFC = lateral prefrontal cortex; Stri = striatum; Error bars: SEM.

632 **Figure 6 Regions showing increased functional connectivity with bilateral striatum during**
633 **third-party punishment decisions (compared with control conditions; A) and plots of**
634 **parameter estimates of PPI in local peak voxel of left IPFC with left (B) /right (C) striatum and**
635 **that of vmPFC with right striatum (D) in four conditions** (i.e. help, help_control, punish,
636 punish_control). PPI = psycho-physiological interaction; IPFC = lateral prefrontal cortex; vmPFC =
637 ventral medial prefrontal cortex; Stri = striatum; Error bars: SEM.

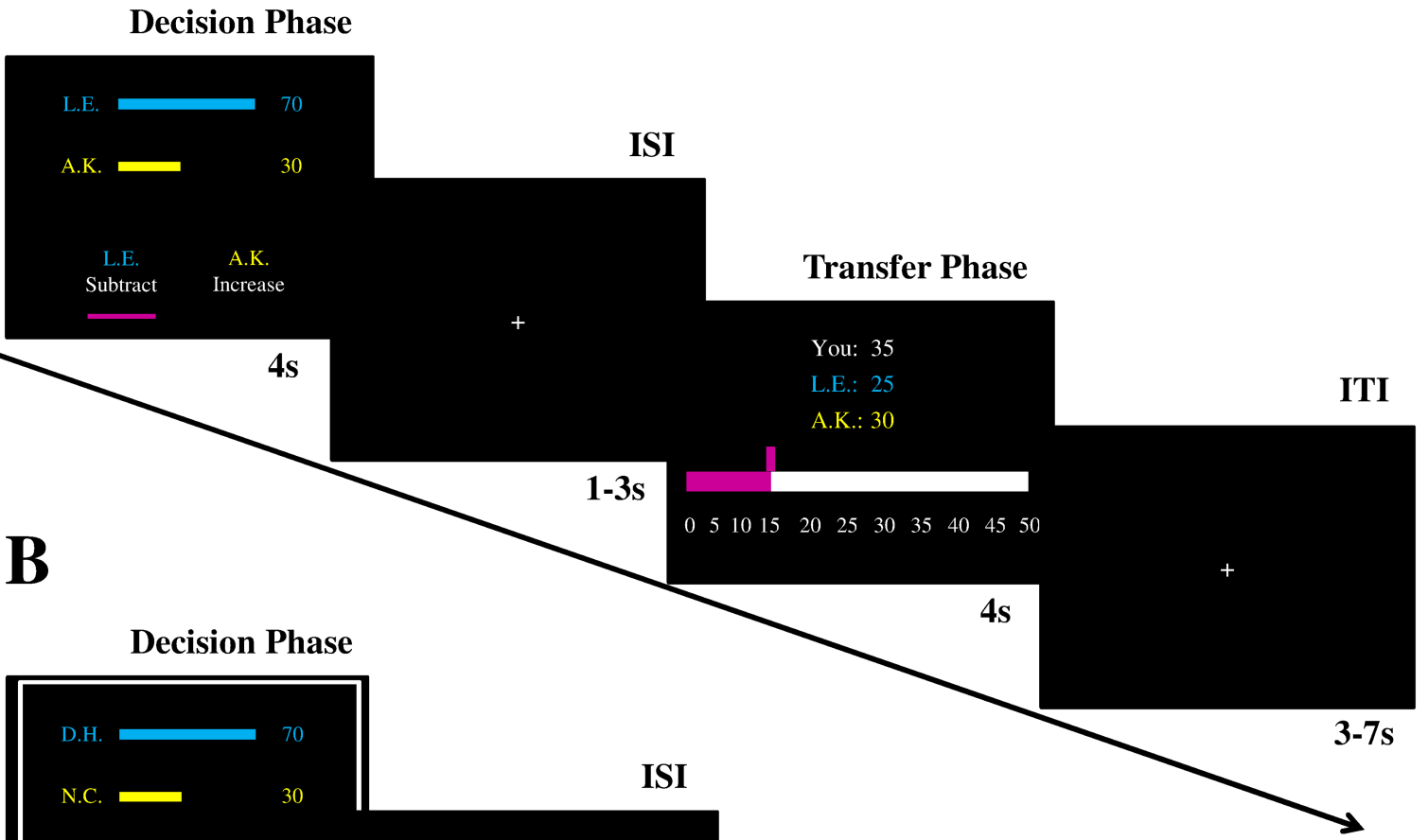
638 8. Supplementary Material

639 The Supplementary Material for this article can be found in a separate document.

640

Figure 1.TIF

A



B

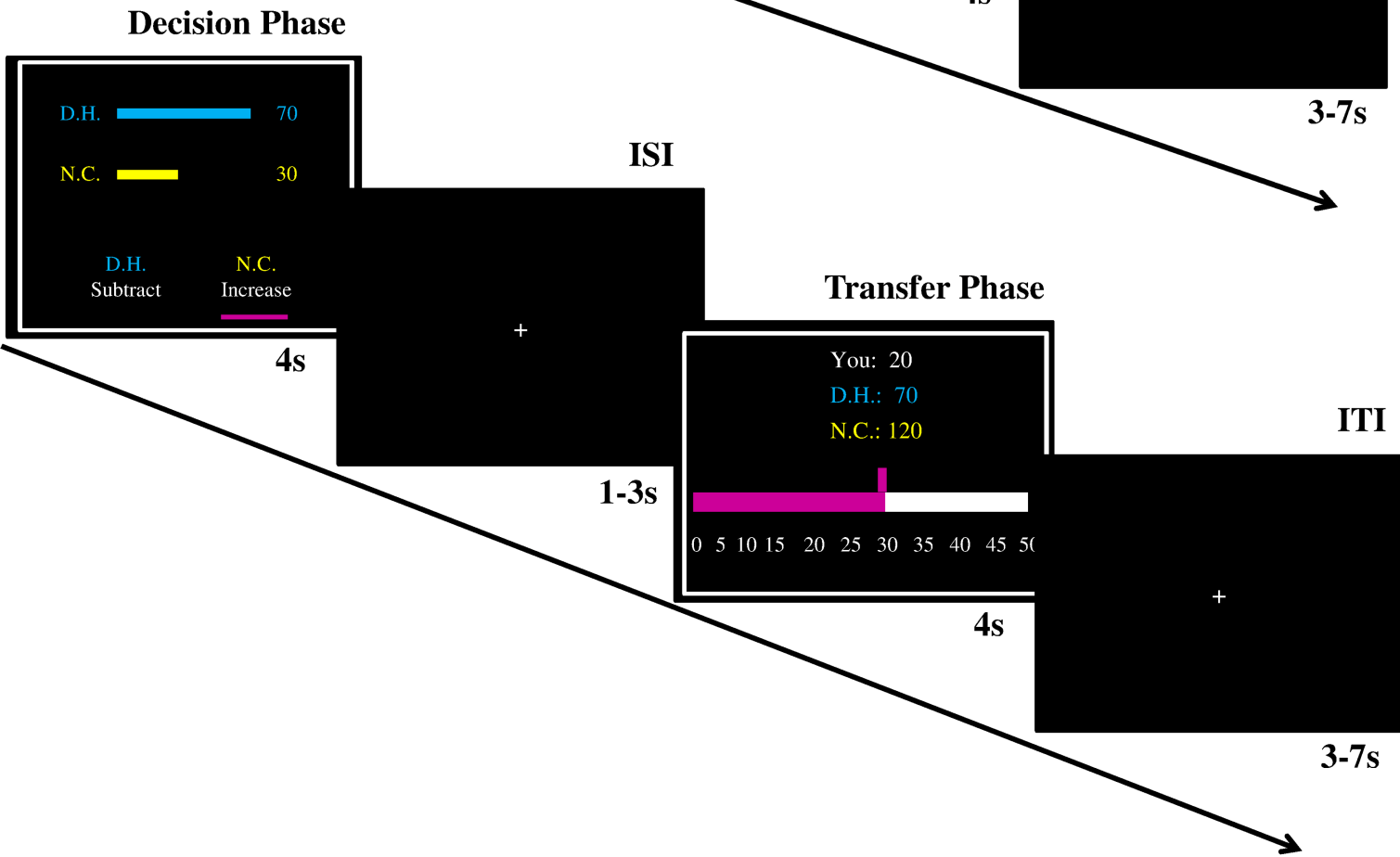


Figure 2.TIF

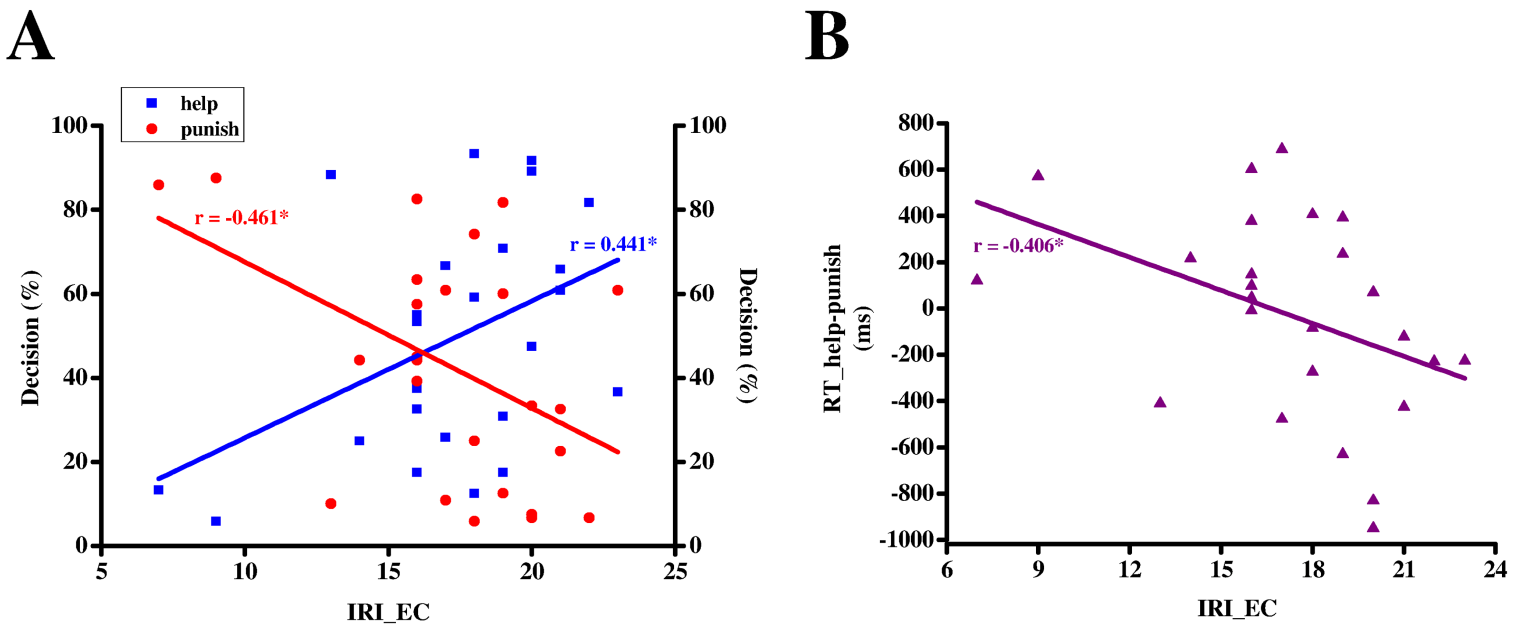
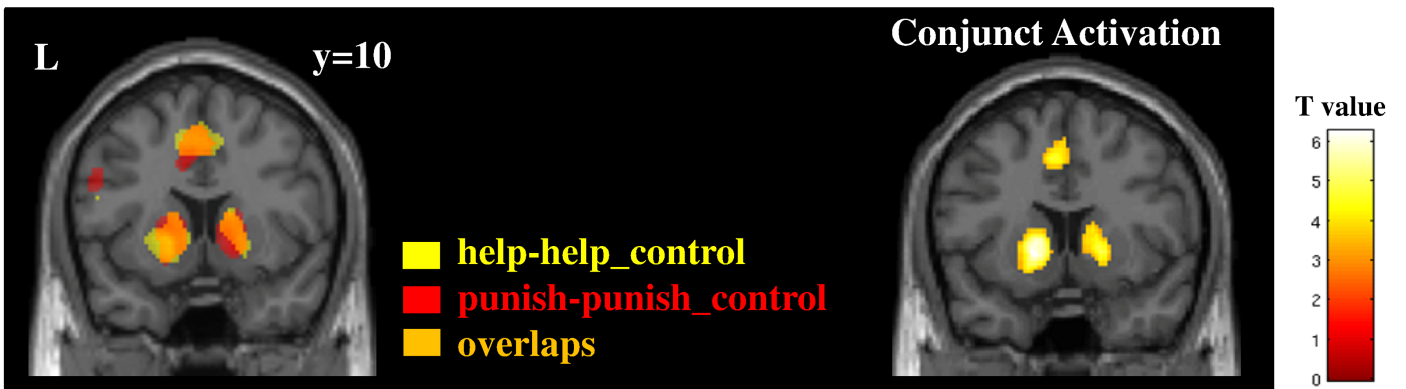
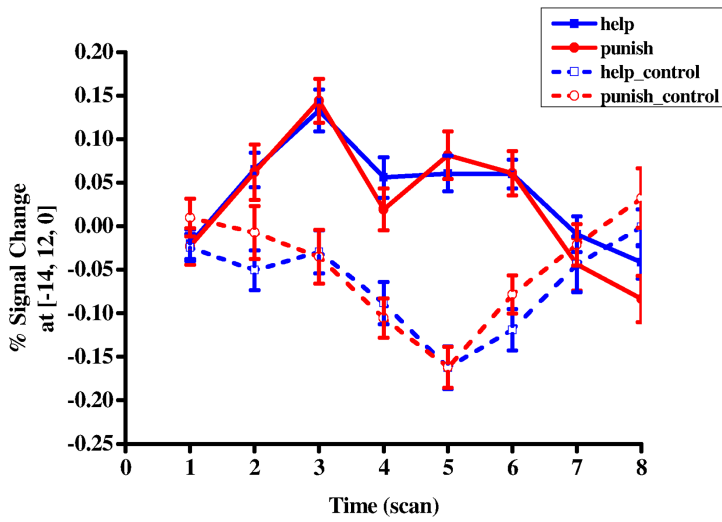


Figure 3.TIF

A



B



C

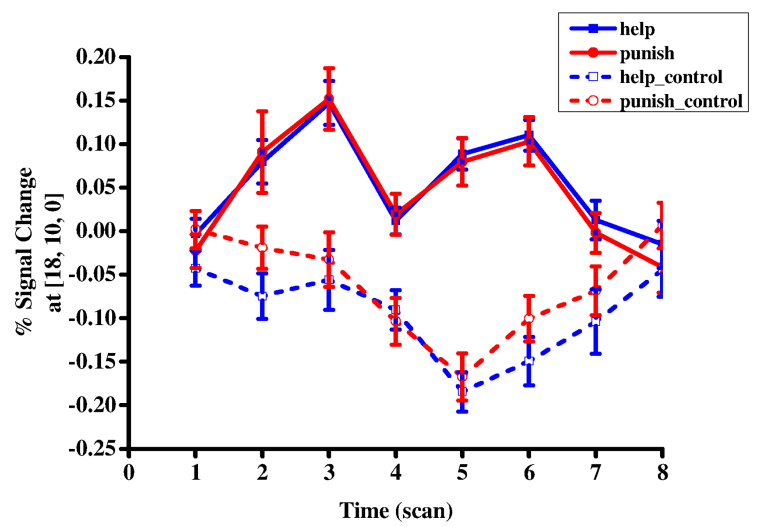
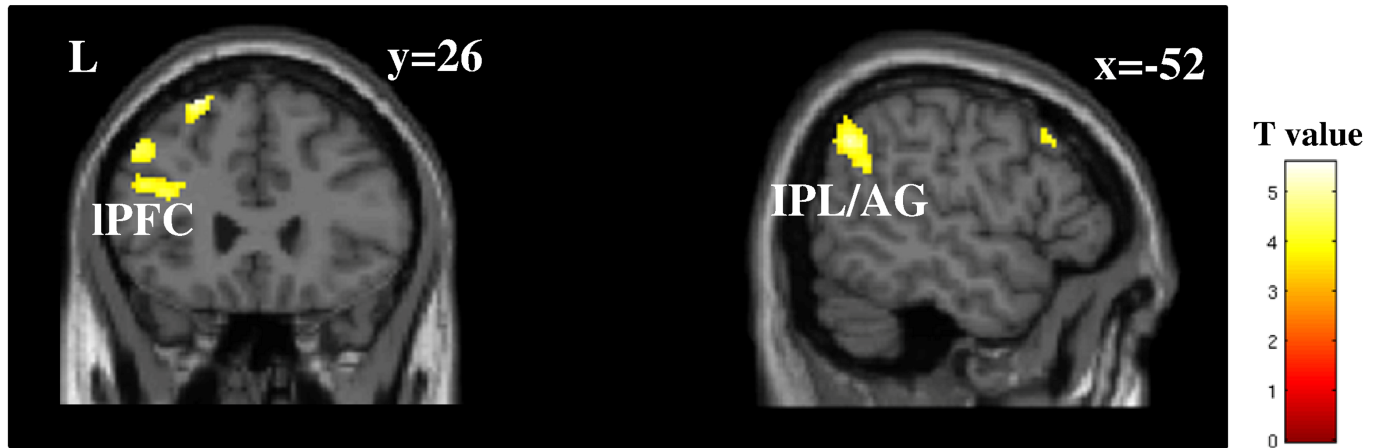
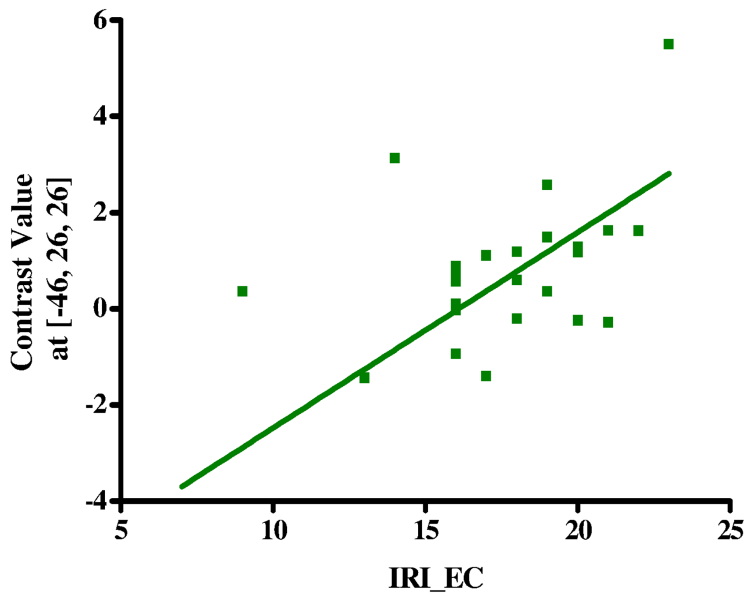


Figure 4.TIF

A



B



C

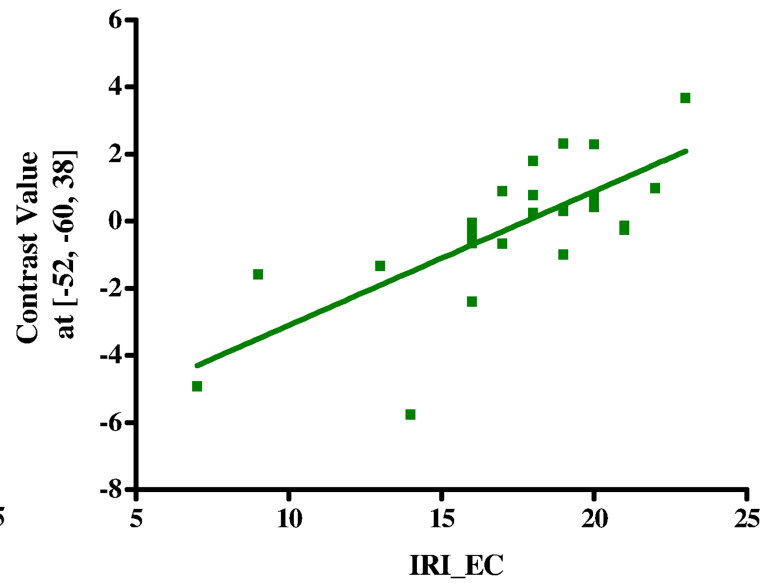
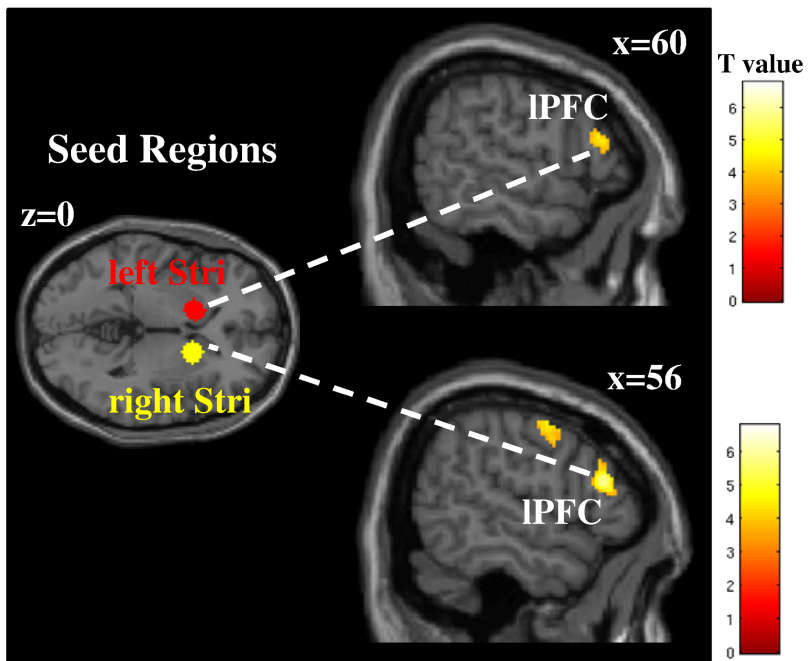
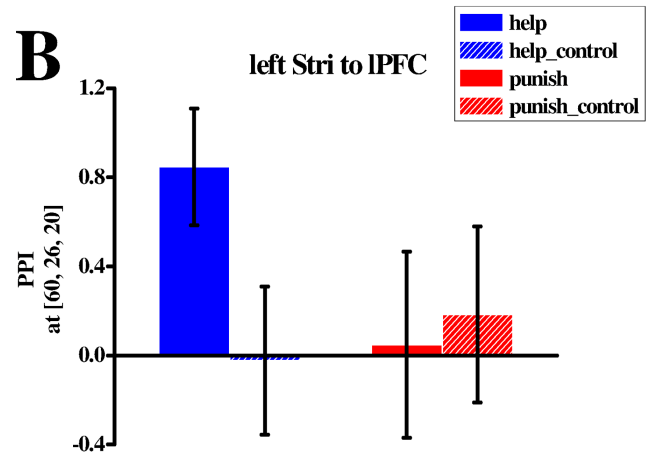


Figure 5.TIF

A



B



C

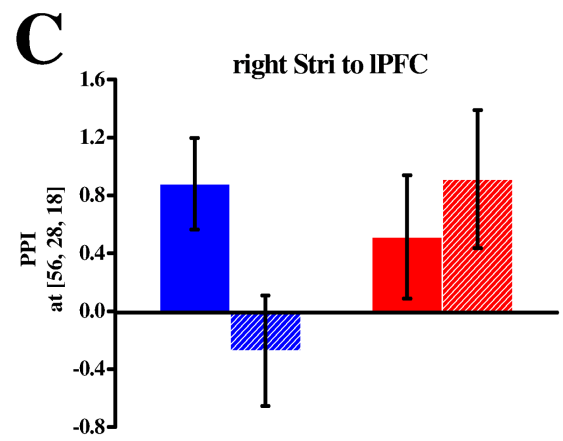


Figure 6.TIF

