
Research Articles: Behavioral/Cognitive

Neural mechanisms underlying individual differences in control-averse behavior

Sarah Rudolf^{1,2}, Katrin Schmelz^{3,4}, Thomas Baumgartner^{1,2}, Roland Wiest⁵, Urs Fischbacher^{3,4} and Daria Knoch^{1,2}

¹Department of Social Psychology and Social Neuroscience, Institute of Psychology, University of Bern, 3012 Bern, Switzerland

²Center for Cognition, Learning and Memory, University of Bern, 3012 Bern, Switzerland

³Department of Economics, University of Konstanz, 78464 Konstanz, Germany

⁴Thurgau Institute of Economics, 8280 Kreuzlingen, Switzerland

⁵Department of Neuroradiology, Inselspital, 3010 Bern, Switzerland

DOI: 10.1523/JNEUROSCI.0047-18.2018

Received: 9 January 2018

Revised: 20 March 2018

Accepted: 25 April 2018

Published: 14 May 2018

Author contributions: S.R. wrote the first draft of the paper; S.R., K.S., T.B., U.F., and D.K. edited the paper; S.R., K.S., T.B., U.F., and D.K. designed research; S.R. and R.W. performed research; S.R. analyzed data; S.R., K.S., T.B., U.F., and D.K. wrote the paper.

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

This project was supported by a grant to DK by the Mens Sana Foundation.

Correspondence to: Sarah Rudolf, Department of Social Psychology and Social Neuroscience, Institute of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland, sarah.rudolf@psy.unibe.ch, or Daria Knoch, Department of Social Psychology and Social Neuroscience, Institute of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland, daria.knoch@psy.unibe.ch

Cite as: J. Neurosci ; 10.1523/JNEUROSCI.0047-18.2018

Alerts: Sign up at www.jneurosci.org/cgi/alerts to receive customized email alerts when the fully formatted version of this article is published.

Accepted manuscripts are peer-reviewed but have not been through the copyediting, formatting, or proofreading process.

Copyright © 2018 Rudolf et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

1 **Neural mechanisms underlying individual differences in**
2 **control-averse behavior**

3
4 *Running title: Neural mechanisms of control-averse behavior*

5
6 Sarah Rudorf^{1,2}, Katrin Schmelz^{3,4}, Thomas Baumgartner^{1,2}, Roland Wiest⁵,
7 Urs Fischbacher^{3,4}, Daria Knoch^{1,2}

8
9 ¹Department of Social Psychology and Social Neuroscience, Institute of Psychology,
10 University of Bern, 3012 Bern, Switzerland

11 ²Center for Cognition, Learning and Memory, University of Bern, 3012 Bern, Switzerland

12 ³Department of Economics, University of Konstanz, 78464 Konstanz, Germany

13 ⁴Thurgau Institute of Economics, 8280 Kreuzlingen, Switzerland

14 ⁵Department of Neuroradiology, Inselspital, 3010 Bern, Switzerland

15
16 Correspondence to:

17 Sarah Rudorf

18 Department of Social Psychology and Social Neuroscience, Institute of
19 Psychology, University of Bern

20 Fabrikstrasse 8, 3012 Bern, Switzerland

21 sarah.rudorf@psy.unibe.ch

22 or

23 Daria Knoch

24 Department of Social Psychology and Social Neuroscience, Institute of
25 Psychology, University of Bern

26 Fabrikstrasse 8, 3012 Bern, Switzerland

27 daria.knoch@psy.unibe.ch

28
29 Number of pages: 47

30 Number of figures: 8

31 Number of tables: 2

32 Number of words in

33 Abstract: 206

34 Introduction: 648

35 Discussion: 1475

36
37 Conflict of Interest: The authors declare no competing financial interests.

38 Acknowledgements: This project was supported by a grant to DK by the Mens Sana
39 Foundation.

40 **ABSTRACT**

41 When another person tries to control one's decisions, some people might comply, but
42 many will feel the urge to act against that control. This control aversion can lead to
43 suboptimal decisions and it affects social interactions in many societal domains. To date,
44 however, it has been unclear what drives individual differences in control-averse
45 behavior. Here, we address this issue by measuring brain activity with fMRI while
46 healthy female and male human participants make choices that are either free or
47 controlled by another person, with real consequences to both interaction partners. In
48 addition, we assessed the participants' affects, social cognitions and motivations via self-
49 reports. Our results indicate that the social cognitions perceived distrust and lack of
50 understanding for the other person play a key role in explaining control aversion at the
51 behavioral level. At the neural level, we find that control-averse behavior can be
52 explained by functional connectivity between the inferior parietal lobule and the
53 dorsolateral prefrontal cortex, brain regions commonly associated with attention
54 reorientation and cognitive control. Further analyses reveal that the individual strength of
55 functional connectivity complements and partially mediates the self-reported social
56 cognitions in explaining individual differences in control-averse behavior. These findings
57 therefore provide valuable contributions to a more comprehensive model of control
58 aversion.

59

60 **SIGNIFICANCE STATEMENT**

61 Control aversion is a prevalent phenomenon in our society. When someone tries to
62 control their decisions, many people tend to act against the control. This can lead to
63 suboptimal decisions, like noncompliance to medical treatments or disobeying the law.
64 The degree to which individuals engage in control-averse behavior, however, varies
65 significantly. Understanding the proximal mechanisms that underlie individual differences
66 in control-averse behavior has potential policy implications, for example when designing
67 policies aimed at increasing compliance with vaccination recommendations, and is
68 therefore a highly relevant research goal. Here, we identify a neural mechanism between
69 parietal and prefrontal brain regions that can explain individual differences in control-
70 averse behavior. This mechanism provides novel insights into control aversion beyond
71 what is accessible through self-reports.

72

73 **INTRODUCTION**

74 When others try to control our decisions, many of us will feel the urge to counteract and
75 thereby reestablish our valued freedom of choice. This aversive reaction to the
76 exogenous control of one's freedom of choice, or in short control aversion, puts a strain
77 on many societal domains, for example in the form of patient noncompliance to
78 psychiatric therapy (De las Cuevas et al., 2014), adolescent defiance against parents
79 (Van Petegem et al., 2015), or employees' reduced work performance when faced with a
80 restrictive employer (Falk and Kosfeld, 2006). Critically, the degree to which individuals
81 engage in control-averse behavior varies largely, which has been documented in
82 numerous studies (Falk and Kosfeld, 2006; Ziegelmeyer et al., 2012; Schmelz and
83 Ziegelmeyer, 2015). What drives these individual differences in control-averse behavior,
84 however, has remained an open question.

85 Previous work has shown that individuals whose decisions are controlled by another
86 person often report thoughts about the other person's motives, such as distrust, and lack
87 of understanding for the other person's decision to control (Falk and Kosfeld, 2006). For
88 example, when an employer requests a minimum effort from her employee, the
89 employee may perceive this as a signal of distrust in her intrinsic work motivation. A
90 separate line of work has highlighted the motivation to restore one's freedom of choice,
91 termed reactance, as the key player in driving control-averse behavior (Brehm, 1966;
92 Miron and Brehm, 2006). For example, the elimination of a choice option can lead to an
93 increased desire for that option, which is interpreted as an indirect strategy of freedom
94 restoration (Miron and Brehm, 2006). Moreover, reactance is assumed to be
95 accompanied by negative affects, such as anger (Dillard and Shen, 2005). Therefore,
96 negative affects and individual tendencies to express one's anger outward might
97 contribute to the display of control-averse behavior. The literature thus delivers several

98 plausible variables that might drive individual control-averse behavior. Much of the
99 support to date, however, comes from post-hoc self-reports or measures of behavioral
100 intentions in hypothetical scenarios. Here, we use a neurophysiological measure of the
101 decision processes during real restrictions of the subjects' freedom of choice. By doing
102 so we aim to identify the proximal mechanisms that give rise to individual differences in
103 control-averse behavior. Specifically, we test whether activation in and functional
104 connectivity with the brain regions that are differentially activated during the restriction of
105 the freedom of choice can explain individual differences in control-averse behavior.
106 Moreover, we investigate to what degree this neurophysiological measure complements
107 and mediates self-report data in predicting individual control-averse behavior.
108 To this end, we combine functional magnetic resonance imaging (fMRI) with a Control
109 aversion task (Fig. 1), in which subjects make decisions that are either free or controlled
110 by another person (Falk and Kosfeld, 2006; Schmelz and Ziegelmeyer, 2015). For each
111 decision, subjects allocate money between themselves and another person by choosing
112 between options that increase in fairness and generosity, called generosity levels.
113 Crucially, the options were designed to establish an intrinsic motivation to choose a high
114 level when subjects can decide freely. When the other person requests a minimum level
115 and thereby tries to control the subject's choice, control-averse behavior is defined as
116 choosing a lower level (Falk and Kosfeld, 2006; Schmelz and Ziegelmeyer, 2015).
117 Therefore, the decrease of average chosen levels when the other person tries to control
118 the subject's decision as opposed to the free decisions serves as a measure of
119 individual control-averse behavior. Critically, the decisions in the task are not
120 hypothetical, but have real consequences for both interaction partners and thus share an
121 important quality with control-averse behavior outside the laboratory. This setup allows
122 us not only to measure control-averse behavior in an ecologically valid fashion, but to

123 investigate the neural responses during the actual decision making process. We find that
124 a neural mechanism involving parietal and prefrontal brain regions complements and
125 partially mediates self-reported social cognition in explaining individual differences in
126 control-averse behavior.

127

128

129 **MATERIALS AND METHODS**

130 **Participants**

131 We recruited a total of 61 students from the University of Bern for participation in this
132 study. Students of economics, psychology and social sciences were excluded from
133 participation to reduce the possibility of prior knowledge of the concept of control
134 aversion. All participants were right-handed, nonsmokers and reported no history of
135 psychological disorders, neurological or cardiovascular diseases. After data acquisition,
136 ten participants were excluded due to excessive movements during fMRI scan (> 5 mm
137 in translation or > 5 degrees in rotation), noncompliance to instructions or technical
138 problems. The remaining 51 participants (23 female; mean 22 ± 3 SD years) were
139 included in the analysis. All participants received a compensation of CHF 50 (\approx USD 50)
140 for participation in the study in addition to the payoff from the Control aversion task
141 described in the next section. The study was approved by the Bern Cantonal Ethics
142 Commission and all participants gave informed, written consent.

143

144 **Experimental design**

145 *Control aversion task*

146 The Control aversion task (Fig. 1) is designed to confront subjects with real restrictions
147 of their freedom of choice by another person and is based on previous work in
148 behavioral economics (Falk and Kosfeld, 2006; Schmelz and Ziegelmeyer, 2015). The
149 gist of the task is that the subject is asked to allocate money between herself and
150 another person, called player A. But before the subject makes a decision, player A can
151 decide to let the subject choose freely (Free condition) or request a minimum amount of
152 money (Controlled condition).

153 For the purpose of this study, subjects were presented with 16 anonymous other
154 persons' (players A's) decisions from a pilot study, in random order. The small number of
155 trials was chosen to increase credibility and reduce possible habituation effects. To
156 ensure equal estimation power of the blood oxygen level-dependent (BOLD) signal
157 across conditions, the players A's decisions were preselected such that the subjects
158 engaged in the same number of trials in the Free and in the Controlled condition, i.e.
159 eight trials per condition. All subjects were informed that the players A's decisions had
160 been prerecorded for logistic reasons, and they were asked to decide as if the respective
161 person was present. To remind subjects of this instruction, we presented the line "A new
162 player A is deciding" for a jittered interval of 2.4-8.6 s at the beginning of each trial.

163 Subjects were also informed that their choices had real consequences in the sense that
164 one trial would be randomly selected and paid out to themselves and the corresponding
165 player A. None of the subjects voiced suspicions about the existence of the players A.
166 After a jittered fixation display of 2-6 s, subjects learned whether the player A let them
167 choose freely (Free condition) or whether the player A requested a minimum amount of
168 monetary units (MUs) (Controlled condition). After a delay of three seconds, subjects

169 made a choice between sets of monetary allocations, called generosity levels, ranging
170 from a selfish (subject : player A, 99:1 MUs) to a more generous, equal allocation (80:80
171 MUs) (all possible generosity levels are depicted in Fig. 1). Subjects made their choice
172 by moving a red selection frame from a random position to their desired option and
173 pressing an OK button. Response times were not constrained to motivate deliberate
174 decisions; subjects were asked, however, to respond as soon as they had come to a
175 decision (response times, mean $5 \pm$ SD 4.3 s). Note, that for the fMRI analysis we
176 separated the times before and after subjects started to move the selection frame to
177 capture the decision window and the motor responses separately. The durations as used
178 in the fMRI analysis are shown in Figure 1. In the Free condition, subjects had the
179 choice between generosity levels one to five (from left to right). In the Controlled
180 condition, subjects' choice was restricted to generosity levels two (97:30 MUs) to five. A
181 central feature of the task is that the player A's payoff increases as a concave function of
182 the generosity levels with relatively small and convex costs for the subject. Moreover,
183 the most generous level (level five) also represented the fairest and equal option and the
184 highest sum of payoffs. These features were added to ensure that subjects are
185 intrinsically motivated to choose a high level, which is a prerequisite for control aversion
186 in this task (Schmelz and Ziegelmeyer, 2015). Lastly, the subject's payoff remains
187 constant for levels two to three. This was done to motivate subjects to choose level three
188 over level two in the Free condition, and to provide space for the choice of a lower level
189 in the Controlled condition that is independent of economic self-interest. The difference
190 between a subject's mean chosen level in the Free condition minus the subject's mean
191 chosen level in the Controlled condition served as the measure of the individual level of
192 control-averse behavior.

193 After another jittered fixation display of 5-8 s, subjects were asked to indicate how they
194 had felt during the decision by rating their unhappiness and anger on 5-point pictorial
195 Self-Assessment Manikins (SAMs) (Bradley and Lang, 1994), each separated by a
196 jittered fixation display of 1-4 s. The unhappiness scale ranged from 1 = “happy” to 5 =
197 “unhappy”, and the anger scale ranged from 1 = “calm” to 5 = “angry”. As a manipulation
198 check we implemented a third scale, the having control scale, which ranged from 1 =
199 “being controlled” to 5 = “having control”. Finally, a fixation cross was displayed for 1.2-
200 6.4 s before the next trial began.

201 Prior to scanning, subjects read the instructions and were quizzed to ensure they had
202 understood the task and its payoff scheme. Subjects then practiced four simulated trials
203 of the Control aversion task outside of the scanner to familiarize themselves with the
204 task timing and the response buttons. Then subjects completed the scanning task in one
205 continuous session of approximately 12 minutes. At the end of the task, one trial was
206 randomly selected for payoff to the subject and the matched player A. Therefore, all
207 trials were incentive compatible to motivate subjects to decide according to their true
208 preferences. The profits in the selected trial were converted into CHF (with 1 MU = CHF
209 0.20 \approx USD 0.20). Based on the task, the subjects received a mean CHF 18.30 \pm
210 1.40 SD, and the players A received a mean CHF 11.10 \pm 3.80 SD.

211

212 *Ratings of perceived distrust, understanding, freedom restoration and fairness*

213 Directly after scanning, we assessed subjects' thoughts during the Control aversion task
214 with a list of items. For each item, subjects were asked to rate how strongly the
215 described thought had influenced their decisions on a 7-point Likert scale ranging from
216 1 = “not at all” to 7 = “absolutely”. Based on the seminal study by Falk and Kosfeld
217 (2006), we assessed subject's perceived distrust and understanding with the items

218 “When player A requests a minimum of generosity, he distrusts me and I dislike that.”
219 (‘perceived distrust’), and “I understand when player A requests a minimum of
220 generosity.” (‘understanding’). Based on reactance theory (Brehm, 1966; Miron and
221 Brehm, 2006), we assessed subjects’ motivation to restore their freedom of choice in the
222 Controlled condition with the item “When player A restricts the generosity levels, I want
223 to use my remaining freedom of choice all the more.” (‘freedom restoration’). In addition,
224 we asked subjects whether fairness had played a role in their own decisions with the
225 item, “I think that my payoff and player A’s payoff should not be too far apart.”
226 (‘fairness’).

227

228 *Assessment of outward directed anger expression*

229 To assess subjects’ general tendency to direct their anger outward, we asked subjects to
230 fill in the German version of the State-Trait Anger Expression Inventory (STAXI)
231 (Spielberger, 1988; Schwenkmezger et al., 1992). The STAXI is comprised of the five
232 subscales state anger, trait anger, inward directed anger expression, outward directed
233 anger expression and controlling one’s anger expression. Here, we focus on the
234 subscale for outward directed anger expression (AO). The AO subscale consists of 8
235 items that describe ways of expressing one’s anger, e.g. “I fly off the handle”. Subjects
236 rated these items on a 4-point Likert scale ranging from 1 = “almost never” to 4 = “almost
237 always”. Based on the subjects’ ratings, the sum scores were computed. In our sample,
238 the AO subscale had an acceptable internal consistency (Cronbach’s alpha = 0.73). On
239 average, subjects had an AO score of mean 12.24 ± SD 3.02 (range 8-22), which is
240 similar to the norm student sample reported in Schwenkmezger et al. (1992).

241

242 **MRI data acquisition and preprocessing**

243 All MRI data were acquired on a Siemens Trio 3.0 Tesla whole-body scanner (Siemens,
244 Erlangen) using a 12-channel head coil. The functional session started off with a
245 localizer scan followed by the Control aversion task implemented in E-Prime 3.0
246 (Psychology Software Tools). The task was projected onto a screen that the subjects
247 viewed through an angled mirror mounted to the head coil. Subjects made their
248 responses on a two-button response box in each hand. While subjects were playing the
249 task, we acquired gradient echo T2*-weighted echo-planar images (EPIs) with BOLD
250 contrast (approx. 400 volumes per subject, 32 slices per volume, ascending order, Field
251 of View 192 x 192 x 110 mm, slice thickness 3 mm, gap 0.45 mm, repetition time 2190
252 ms, echo time 30 ms, flip angle 90°). Volumes were acquired in axial orientation at a
253 +15° tilt to the anterior commissure-posterior commissure line. After the functional
254 session, T1-weighted 3D modified driven equilibrium Fourier transformation (MDEFT)
255 images were acquired from each subject (176 slices, Field of View 256 x 224 x 176 mm,
256 slice thickness 1 mm, no gap, repetition time 7.92 ms, echo time 2.48 ms, flip angle 16°).
257 Preprocessing of the functional images was implemented in the MATLAB based
258 software Statistical Parametric Mapping 12 (SPM12, version r6685;
259 <http://www.fil.ion.ucl.ac.uk/spm>). Preprocessing included motion correction (realignment
260 to the mean EPI), segmentation of the T1 image into six tissue classifications (gray
261 matter, white matter, cerebro-spinal fluid, bone, soft tissue and air tissue), application of
262 this segmentation to the mean EPI, coregistration of all EPIs to the mean EPI using the
263 pullback procedure in the SPM12 deformation tool and normalization of all EPIs to MNI
264 standard space (Montreal Neurological Institute, <http://www.bic.mni.mcgill.ca>) (Evans et
265 al., 1993). Finally, we smoothed the EPIs with a 4 mm full width at half maximum
266 Gaussian kernel.

267

268 **Analysis aim and structure**

269 The central aim of our analyses was to identify a neurophysiological mechanism that can
270 explain individual differences in control-averse behavior in addition to or beyond self-
271 report data. To this end, our analyses followed a hierarchical structure. First, we
272 identified the best predictor of individual control-averse behavior based on self-report
273 data. Second, we identified a neurophysiological mechanism that predicts individual
274 control-averse behavior. Third, we identified the best combination of predictors based on
275 both self-report and neural data. Fourth, we tested whether the neural predictor
276 mediates the self-report data in predicting individual control-averse behavior.

277

278 **Behavioral data analyses**

279 All behavioral data were analyzed using the MATLAB Statistics and Machine Learning
280 Toolbox (R2015b, MathWorks). Because the behavioral data did not follow normal
281 distributions as assessed by Kolmogorov-Smirnov tests, nonparametric tests were
282 applied. Paired samples were compared using the Wilcoxon signed rank test.
283 Correlations were assessed using Spearman's rho as well as bisquare robust
284 regressions. For all behavioral analyses, two-tailed p values are reported.

285

286 *Identifying the best predictor of individual control-averse behavior based on self-report*
287 *data*

288 We first identified the best predictor of individual control-averse behavior based on self-
289 report data. To this end, we ran a series of generalized linear models using the function
290 fitglm as implemented in the MATLAB Statistics and Machine Learning Toolbox
291 (R2015b, MathWorks). For each model, the dependent variable was the individual level

292 of control-averse behavior, as measured by the mean chosen level in the Free condition
293 minus the mean chosen level in the Controlled condition. The self-report variables
294 served as predictors. For conciseness, we report only models with predictors that
295 showed a significant correlation with individual control-averse behavior. To reduce
296 multicollinearity among the predictors, we computed two new variables using principal
297 component analysis as implemented in the MATLAB function `pca`. The new variable
298 'social cognition' is the first principal component of the normalized ratings of the item
299 'perceived distrust' (coefficient 0.88) and the reversed item 'understanding' (coefficient
300 0.48). The second new variable 'negative affect' is the first principal component of the
301 normalized mean unhappiness rating (coefficient 0.80) and the normalized mean anger
302 rating in the Controlled minus the Free condition (coefficient 0.59). As predictors we
303 used combinations of main effects and interactions of 'social cognition', 'negative affect',
304 and the normalized ratings of the item 'freedom restoration'. The most relevant models
305 are illustrated in Figure 4. We compared the models using the Bayesian information
306 criterion (BIC) and R^2 to identify the best model fit. Lower values in BIC and greater
307 values in R^2 indicate better model fits.

308

309 **fMRI data analyses**

310 The statistical analysis of the fMRI data was also carried out in SPM12 (version r6685).
311 We modeled each subject's BOLD response with a General Linear Model (GLM) that
312 was estimated using SPM12's standard hemodynamic response function and a high
313 pass filter of 128 Hz as well as correction for intrinsic autocorrelations. SPM12's internal
314 masking threshold for the estimation of the beta parameters was set to 0.4 to ensure
315 inclusion of subcortical brain regions. The GLM contained two regressors of interest as
316 boxcar functions: (1) decisions in the Controlled condition, and (2) decisions in the Free

317 condition (each with a duration from the respective onset of the choice options until the
318 first button press, illustrated as ‘decision window’ in Fig. 1). Note that due to a high
319 consistency in the subjects’ choices and therefore in the subjects’ and player A’s payoff
320 within each condition and subject (Fig. 2B), it was not feasible to additionally control for
321 the subjects’ or player A’s payoff in the GLM. As nuisance regressors, we modeled (3)
322 the display of the text “A new player A is deciding...” (duration 2.4-8.6 s), (4) motor
323 response (duration from the first button press until press of the OK button), (5)
324 unhappiness rating (duration = reaction times), (6) anger rating (duration = reaction
325 times), (7) manipulation check, i.e. feeling of being controlled rating (duration = reaction
326 times), (8) six motion parameters. For every subject we created contrast images for the
327 two regressors of interest.

328 At the group level, we used random effects analyses. For all random effects analyses,
329 we applied whole-brain correction for multiple comparisons at the cluster level: We
330 calculated the corrected cluster extent (k_E) for each t test using Gaussian Random-field
331 theory as implemented in SPM12 with a cluster-defining individual voxel threshold of
332 $t = 2.68$ ($p < 0.005$) to achieve a FWE-corrected statistical threshold of $p_{FWE} < 0.05$
333 (minimum $k_E > 40$, range 40-44).

334 The aim of the fMRI analysis was to identify a neurophysiological mechanism that can
335 predict individual differences in control-averse behavior. Specifically, we investigate
336 whether activations in and interactions with the brain regions that are differentially
337 activated for decisions in the Controlled and the Free condition correlate with individual
338 control-averse behavior. We did so in three fMRI analysis steps, which will be described
339 in the following sections.

340

341 *fMRI analysis step 1: Localization of brain regions differentially activated for decisions in*
342 *the Controlled and the Free condition*

343 To identify the brain regions that are differentially activated during decisions in the
344 Controlled and the Free condition, we tested the corresponding contrast images in a
345 paired t test at the group level. Because we had no strong anatomical hypotheses, we
346 applied whole-brain corrected analysis. Based on the paired t test, we created two
347 masks for all suprathreshold voxels within a 10-mm sphere around the group peak voxel
348 in the right and left inferior parietal lobule (IPL), respectively, at a threshold of $p < 0.005$,
349 uncorrected (peak MNI coordinates for right IPL: 39 -40 40; for left IPL: -42 -40 47,
350 illustrated in Fig. 5). The spheres were applied to isolate the activation in the IPL from
351 more posterior activation. The masks were used to extract and illustrate the mean beta
352 estimates as implemented in the MarsBaR toolbox (Brett et al., 2002), as well as for
353 search volumes in the functional connectivity analyses and time course analyses that will
354 be described in the fMRI analysis step 3.

355

356 *fMRI analysis step 2: Covariate analysis of activation differences for decisions in the*
357 *Controlled and the Free condition and control-averse behavior*

358 The second step of the fMRI analysis was to investigate whether individual control-
359 averse behavior could be predicted by activation differences for decisions in the
360 Controlled and Free condition. To test this, we included the individual level of control-
361 averse behavior as a covariate in the paired t test (random effects analysis), using a
362 whole-brain analysis. The individual level of control-averse behavior was computed as
363 the mean chosen level in the Free condition minus the mean chosen level in the
364 Controlled condition, with the result that increasing values reflect increasing levels of
365 control-averse behavior.

366

367 *fMRI analysis step 3: Covariate analysis of the functional connectivity seeded in the IPL*
368 *and control-averse behavior*

369 The third step of the fMRI analysis was to investigate whether individual control-averse
370 behavior could be explained by neural interactions with the brain regions that are
371 differentially active for decisions in the Controlled and Free condition. For this purpose
372 we conducted functional connectivity analyses seeded in the right and left IPL as
373 identified in the paired t test for decisions in the Controlled > Free condition. To assess
374 the functional connectivity, we used psychophysiological interaction (PPI) analysis with
375 two psychological factors of interest that were derived from the GLM: (1) decisions in the
376 Controlled condition; (2) decisions in the Free condition. We extracted single-subject
377 time courses in the right and the left IPL, respectively, as follows: using the search
378 volumes derived from the paired t test for decisions in the Controlled > Free condition at
379 the group level (illustrated in Fig. 6), we identified, for each subject, the peak Z value for
380 the contrast of decisions in the Controlled > Free condition and extracted the first BOLD
381 signal eigenvariate from a 5-mm sphere around this individual peak. This approach was
382 chosen to account for between-subject variability in the spatial location of the peak
383 activation. The extracted BOLD signal eigenvariate was then deconvolved and multiplied
384 with the two psychological factors of interest to create the PPI terms (Controlled PPI,
385 Free PPI), which were then convolved with the standard SPM12 hemodynamic response
386 function. Lastly, for each seed, the two PPI terms, the BOLD signal eigenvariate, and all
387 regressors described in the GLM were entered into a new GLM (GLM-PPI). For all
388 subjects, we created contrast images for the two PPI terms. To identify brain regions that
389 show an increased functional connectivity with the right and left IPL, respectively, we
390 tested the associated contrast images Controlled PPI > Free PPI in two separate paired t

391 tests at the group level (random effects analyses). Finally, to test whether the functional
392 connectivity seeded in the IPL predicts control-averse behavior, we included the
393 individual level of control-averse behavior as a covariate in the paired t tests of
394 Controlled PPI > Free PPI (random effects analyses), using whole-brain analyses.
395 Based on the covariate analysis, we created two new masks for all suprathreshold
396 voxels in the right and left dorsolateral prefrontal cortex (dIPFC)/ middle frontal gyrus,
397 respectively, at a threshold of $p < 0.005$, uncorrected (Table 1, Fig. 6). These masks
398 were used to extract and illustrate the mean beta estimates as implemented in the
399 MarsBaR toolbox (Brett et al., 2002) (Fig. 5), and as search volumes for additional time
400 course analyses (Fig. 6) as follows.

401 To further examine individual differences in the temporal characteristics of the BOLD
402 signal underlying the decisions in the Controlled and Free condition in the seed (bilateral
403 IPL) and target regions (bilateral dIPFC/ middle frontal gyrus) of the functional
404 connectivity analysis, we performed *post hoc* time course analyses using the search
405 volumes as described above. For each subject and each search volume, we identified
406 the peak Z value for the contrast of decisions in the Controlled > Free condition and
407 extracted the raw event-related BOLD response from a 5-mm sphere around this
408 individual peak, which was identical to the procedure used in the PPI analysis. Event-
409 related BOLD responses were estimated by two Finite Impulse Response models for
410 decisions in the Controlled condition and decisions in the Free condition, respectively,
411 adjusted for nuisance effects of the motion regressors and resampled to time bins of
412 0.5 s as implemented in the rfxplot toolbox (Gläscher, 2009). We then divided the
413 subjects into groups of not control-averse subjects (with levels of control-averse
414 behavior ≤ 0 , $n = 10$) and control-averse subjects (with levels of control-averse behavior
415 > 0 , $n = 41$) and plotted the averaged time courses across subjects in each group

416 separately for decisions in the Controlled and the Free condition (Fig. 6). Note, that the
417 raw event-related BOLD signal is independent of any model assumptions. The time
418 course analyses therefore provide additional insights into the temporal characteristics of
419 the BOLD signal in the target regions. Due to the use of non-independent masks,
420 however, it is important to note that the time course analyses were not used to infer the
421 magnitude of the effect Controlled > Free condition.

422

423 **Identifying the best combination of predictors of individual control-averse**
424 **behavior based on self-report and neural data**

425 Building upon the behavioral results and the result of the functional connectivity analysis,
426 we next investigated whether models based on self-report data could be improved by
427 including neural data. To this end, we ran a new series of generalized linear models
428 using the function fitglm as implemented in the MATLAB Statistics and Machine
429 Learning Toolbox (R2015b, MathWorks). For each model, the dependent variable was
430 the individual level of control-averse behavior, as measured by the mean chosen level in
431 the Free condition minus the mean chosen level in the Controlled condition.

432 We compared the best model based on self-report data with models based on the neural
433 data and combinations of neural and self-report data. As neural predictor we used the
434 difference between the subject-wise estimate of the connectivity between right IPL and
435 right dlPFC during decisions in the Controlled and the Free condition (Controlled PPI –
436 Free PPI). This neural predictor was combined with main effects of and interactions with
437 the predictors ‘social cognition’, ‘negative affect’, and ‘freedom restoration’. The most
438 relevant models are illustrated in Figure 7. Again, we compared the models with regard
439 to the BIC and R^2 .

440

441 **Mediation analysis of self-report and neural predictors of individual control-averse**
442 **behavior**

443 Building upon the result of the model comparisons, we next investigated the association
444 between social cognition, the right IPL-dIPFC connectivity and control-averse behavior.
445 To this end, we performed a mediation analysis using the MATLAB based mediation
446 toolbox described by Wager et al. (2008, <https://github.com/canlab/MediationToolbox>).
447 We based the test on three criteria, which are illustrated in the three-variable path model
448 in Figure 8. First, the predictor must be related to the mediating variable (path a).
449 Second, the mediator must be related to the outcome after controlling for the predictor
450 (path b). Third, the mediation effect defined as product of the a and b path coefficients
451 ($a*b$) must be significant. A significant mediation effect indicates that the mediator
452 significantly reduces and therefore explains the predictor-outcome relationship
453 (difference between path c and c'). If the predictor still explains significant variance in the
454 outcome after controlling for the mediator (path c'), we speak of a partial mediation.
455 A mediation analysis is conceptually different from a moderation analysis (see model 10
456 in Fig. 7), which tests whether the level of the moderating variable can predict the
457 strength of the relationship between the predictor and the outcome (Baron and Kenny,
458 1986; Wager et al., 2008). In other words, a moderator indicates *when* a predictor-
459 outcome association occurs, whereas a mediator explains *how or why* such an effect
460 occurs (Baron and Kenny, 1986). We therefore ran the mediation analysis to test
461 whether the right IPL-dIPFC connectivity represents the mechanism through which social
462 cognition affects control-averse behavior.
463 As the predictor we used the subject-specific variable 'social cognition'. The mediator
464 was the difference between the subject-wise estimate of the connectivity between right
465 IPL and right dIPFC during decisions in the Controlled and the Free condition (Controlled

466 PPI – Free PPI). The outcome was the individual level of control-averse behavior, as
467 measured by the mean chosen level in the Free condition minus the mean chosen level
468 in the Controlled condition. Statistical significance was assessed using a bootstrap test
469 with 1000 samples.

470

471

472 **RESULTS**

473 **Behavioral results**

474 *Control-averse behavior and its association with negative affect, perceived distrust,*
475 *understanding, and freedom restoration*

476 While lying in the fMRI scanner, subjects made choices under two conditions (Fig. 1): In
477 the Free condition, subjects could choose freely among five allocation options, called
478 generosity levels, ranging from selfish to more generous and equal monetary allocations
479 between themselves and another person. In the Controlled condition, the other person
480 requested a minimum of level two and thereby eliminated the most selfish and unequal
481 option. A manipulation check showed that subjects indeed indicated having more control
482 in the Free condition (mean $4.42 \pm$ SD 0.73, median 4.75) than in the Controlled
483 condition (mean $3.88 \pm$ SD 0.88, median 4.00; Wilcoxon signed rank test, two-tailed,
484 $Z = 4.69$, $p < 0.001$, Hodges-Lehmann estimator of differences 0.63, 95 % CI [0.38,
485 0.94], Fig. 2A).

486 First, we tested whether the restriction of the freedom of choice had an effect on
487 subjects' generosity as measured by the chosen generosity level. As expected, subjects
488 chose, on average, lower generosity levels in the Controlled condition (mean $3.50 \pm$ 0.78
489 SD, median 3.50) than in the Free condition (mean $4.34 \pm$ 0.57 SD, median 4.50;
490 Wilcoxon signed rank test, two-tailed, $Z = -5.64$, $p < 0.001$, Hodges-Lehmann estimator

491 of differences -1.00 , 95 % CI $[-1.19, -0.81]$, Fig. 2B). Note that the statistical test was
492 corrected for a bottom effect, following the procedure by Falk and Kosfeld (2006).
493 Subjects demonstrated high consistency in their choice preferences: they showed a
494 variance of mean 0.31 ± 0.33 SD, median 0.21 in the Controlled condition and a
495 variance of mean 0.33 ± 0.37 SD, median 0.21 in the Free condition (Fig. 2C). We
496 therefore averaged each subjects' choices within each condition and used the difference
497 between each subject's mean chosen level in the Free condition minus the subject's
498 mean chosen level in the Controlled condition as the measure of the individual level of
499 control-averse behavior. Notably, the individual levels of control-averse behavior varied
500 from -0.25 to 2.13 (mean 0.82 ± 0.64 SD, median 0.88), a variation that stems mostly
501 from the mean chosen levels in the Controlled condition rather than the Free condition
502 as illustrated in in Figure 2B-C. In other words, subjects chose similarly high levels in the
503 Free condition, whereas choices are more heterogeneous in the Controlled condition.
504 For two subjects, the level of control-averse behavior was -0.25 , which did not result
505 from systematic choices, but rather from a single outlier choice of a lower level in the
506 Free condition. Because these subjects otherwise demonstrated zero difference in their
507 choices between the two conditions, they were treated as not control-averse.
508 Second, we tested whether subjects' individual control-averse behavior was associated
509 with negative affects (Dillard and Shen, 2005). To capture negative affects, we used
510 trial-by-trial anger ratings of unhappiness and anger on pictorial 5-point SAM scales
511 (Bradley and Lang, 1994). Indeed, we found a significant association of control-averse
512 behavior with both negative affect ratings: the unhappier (Spearman's $\rho = 0.49$,
513 $p < 0.001$; robust $R^2 = 0.26$, $p < 0.001$) and the angrier (Spearman's $\rho = 0.46$,
514 $p = 0.001$; robust $R^2 = 0.23$, $p < 0.001$) subjects were in the Controlled compared with
515 the Free condition, the greater was their individual level of control-averse behavior (Fig.

516 **3A**). To additionally assess trait anger expression, we used a task-independent anger
517 expression inventory (STAXI, Schwenkmezger et al., 1992). Subjects' general tendency
518 to direct anger expression outward, however, did not correlate significantly with the
519 individual level of control-averse behavior (Spearman's $\rho = -0.01$, robust $R^2 < 0.01$,
520 both $p > 0.9$, Fig. **3A**). Other subscales of the STAXI also showed no significant
521 association with control-averse behavior.

522 Third, we tested the association between subjects' individual control-averse behavior
523 and their self-reported thoughts as assessed by ratings after scanning. For each rating,
524 subjects were asked to indicate how strongly the described thought had influenced their
525 decision in the Control aversion task. Consistent with previous work (Falk and Kosfeld,
526 2006), we found that subjects demonstrated more control-averse behavior the more they
527 perceived the choice restriction as a signal of distrust by the other person (Spearman's
528 $\rho = 0.60$, robust $R^2 = 0.32$, both $p < 0.001$, Fig. **3B**). By contrast, subjects
529 demonstrated less control-averse behavior the higher they rated understanding the other
530 person's request in the Controlled condition (Spearman's $\rho = -0.66$, robust $R^2 = 0.37$,
531 both $p < 0.001$). Then we tested whether the motivation for freedom restoration had
532 influenced the subjects' decisions. Consistent with reactance theory (Brehm, 1966;
533 Miron and Brehm, 2006), our subjects' self-reported motivation to use their remaining
534 freedom of choice correlated significantly and positively with their level of control-averse
535 behavior (Spearman's $\rho = 0.37$, $p = 0.008$, robust $R^2 = 0.17$, $p = 0.003$, Fig. **3B**).

536 Lastly, we asked subjects whether fairness had played a role in their decisions, i.e. the
537 thought that their own payoff and the other person's payoff should not be too far apart.
538 Interestingly, fairness correlated positively with the average chosen level within both the
539 Controlled condition (Spearman's $\rho = 0.51$, robust $R^2 = 0.28$, both $p < 0.001$) and the
540 Free condition (Spearman's $\rho = 0.48$, robust $R^2 = 0.26$, both $p < 0.001$), but was not

541 significantly associated with control-averse behavior (Spearman's rho = -0.20,
542 $p = 0.163$, robust $R^2 = 0.04$, $p = 0.144$, Fig. 3B).

543

544 *Social cognition is the best self-report predictor of individual control-averse behavior*

545 Next, we aimed to identify the best predictor of individual control-averse behavior based
546 on self-report data. To this end, we computed and compared a series of generalized
547 linear models. As predictors we focused on the self-reported variables that showed a
548 significant correlation with control-averse behavior (Fig. 3). To reduce multicollinearity
549 among the predictors, we applied principal component analyses and computed the new
550 variables 'social cognition' and 'negative affect'. The normalized ratings of the item
551 'freedom restoration' served as a third predictor. Model comparisons revealed that,
552 based on the self-report data, the following model had the best model fit (Fig. 4, Table
553 2):

$$y_i = \beta_0 + \beta_1 \text{SocialCognition}_i + \varepsilon_i$$

554 where y_i is the level of control-averse behavior for subject i , and *SocialCognition* is the
555 first principal component of the normalized ratings of the items 'perceived distrust' and
556 the reversed item 'understanding'. This model performed better in predicting individual
557 control-averse behavior than any model that included negative affect or the motivation
558 for freedom restoration either as main effects or interaction terms.

559

560 **Neuroimaging results**

561 *Control-averse behavior is predicted by neural interactions between the right IPL and the*
562 *dorsolateral prefrontal cortex (dlPFC)*

563 The aim of the fMRI analysis was to identify a neurophysiological mechanism that can
564 predict control-averse behavior. Specifically, we aimed to test whether neural responses
565 and their interactions could explain individual differences in control-averse behavior. To
566 do this, we ran covariate analyses between the individual control-averse behavior and
567 neural activity in the brain regions that are differentially activated during decisions in the
568 Controlled and the Free condition, as well as the functional connectivity seeded in these
569 brain regions.

570 In a first step, the brain regions that are more strongly activated during decisions in the
571 Controlled than in the Free condition were localized. We estimated a GLM that models
572 the BOLD responses for decisions in the Controlled and the Free condition, respectively.
573 The respective single-subject contrast images were then compared in a paired t test. We
574 found that the right IPL (peak MNI coordinates 39 -40 40, $t = 3.99$, $p_{FWE} < 0.001$, whole-
575 brain family-wise error (FWE)-corrected at the cluster level), the left IPL (peak MNI
576 coordinates -42 -40 47, $t = 3.76$, $p_{FWE} = 0.042$), clusters in the bilateral superior parietal
577 lobule extending into the occipital cortex (peak MNI coordinates right 15 -73 57, $t = 4.42$,
578 $p_{FWE} < 0.001$; left -21 -64 43, $t = 4.43$, $p_{FWE} < 0.001$) and the right occipital cortex (peak
579 MNI coordinates 39 -79 33, $t = 4.01$, $p_{FWE} = 0.042$) were more strongly activated during
580 decisions in the Controlled than in the Free condition.

581 In a second step, we tested whether these activation differences between decisions in
582 the Controlled and in the Free condition could explain individual differences in control-
583 averse behavior by including the individual level of control-averse behavior as a
584 covariate in the paired t test of the contrast images for decisions in the Controlled and

585 the Free condition. This covariate analysis revealed no significant association between
586 control-averse behavior and the activation differences between decisions in the
587 Controlled and the Free condition, even at a more liberal statistical threshold of
588 $p < 0.005$, uncorrected.

589 In a third step, we asked whether individual differences in control-averse behavior could
590 instead be explained by functional connectivity patterns. As the seed region of the
591 functional connectivity, we focused on the bilateral IPL due to its suggested role in
592 subjective choice restrictions (Filevich et al., 2013) and attention reorientation (Corbetta
593 et al., 2008). Accordingly, the above described peak activation clusters in the bilateral
594 IPL were used as search volumes for individual subjects' seeds for the functional
595 connectivity analyses (illustrated in Fig. 5 and 6). To assess the functional connectivity,
596 we performed two psychophysiological interaction (PPI) analyses that included separate
597 interaction terms between the right and left IPL BOLD time series, respectively, and
598 regressors indicating decisions in the Controlled and the Free condition (Controlled PPI,
599 Free PPI). We searched for brain regions whose functional connectivity with the IPL
600 predicted control-averse behavior by including the individual level of control-averse
601 behavior as a covariate in the paired t test of the contrast images for Controlled PPI >
602 Free PPI. Whereas the covariate analysis seeded in the left IPL revealed no significant
603 results, we found that for Controlled PPI > Free PPI, the right IPL showed increased
604 functional coupling with the right dorsolateral prefrontal cortex (dlPFC)/ middle frontal
605 gyrus ($p_{FWE} < 0.001$), the left angular gyrus ($p_{FWE} < 0.001$), the right precuneus
606 ($p_{FWE} = 0.047$), the left dlPFC ($p_{FWE} = 0.042$) and the left IPL ($p_{FWE} = 0.033$) as a function
607 of control-averse behavior (Fig. 5, Table 1). No significant negative association was
608 observed. Complementary PPI analyses seeded in the superior parietal lobule and the
609 occipital cortex revealed no significant association with control-averse behavior. To

610 inspect whether the positive correlation was driven by either one of the conditions, we
611 extracted the mean beta estimates across the functional clusters of the bilateral dIPFC
612 for the Controlled PPI and the Free PPI regressor, separately, and plotted them against
613 the individual level of control-averse behavior (Fig. 5). This inspection revealed that right
614 IPL-dIPFC connectivity during the decisions increased with control-averse behavior in
615 the Controlled condition and decreased with control-averse behavior in the Free
616 condition. Hence, the higher the individual level of control-averse behavior, the greater is
617 the change in right IPL-dIPFC connectivity during decisions in the Controlled in contrast
618 to the Free condition. In addition, time course analyses showed that activation in the
619 bilateral IPL increases immediately after the onset of the choice options, irrespective of
620 individual control-averse behavior (Fig. 6). In contrast, activation in the bilateral dIPFC
621 synchronizes with activation in the IPL only for control-averse subjects and only during
622 decisions in the Controlled condition.

623

624 *The connectivity between right IPL and dIPFC complements self-reported social*
625 *cognition in predicting individual control-averse behavior*

626 Next, we aimed to identify the best combination of predictors of control-averse behavior
627 based on both self-report and neural data. Specifically, we tested whether the functional
628 connectivity with the IPL complements or exceeds the self-reports in predicting control-
629 averse behavior. To this end, we computed a set of new generalized linear models that
630 included the neural data. As neural predictor, 'PPI', we used the subject-wise beta
631 estimate of the Controlled PPI minus the Free PPI regressor between the right IPL and
632 the right dIPFC. We focused on the connectivity of the right IPL with the dIPFC because
633 of their frequent coactivation during attention reorientation (Corbetta et al., 2008) and
634 context-dependent decision making (Daw et al., 2006; Boorman et al., 2009; Rudolf and

635 Hare, 2014). This neural predictor was combined with main effects of and interactions
636 with the predictors 'social cognition', 'negative affect', and 'freedom restoration'. Model
637 comparisons revealed that a model that combined main effects of 'social cognition' and
638 'PPI' had the best overall model fit (model 9, Fig. 7, Table 2):

$$y_i = \beta_0 + \beta_1 \text{SocialCognition}_i + \beta_2 \text{PPI}_i + \varepsilon_i$$

639 This model performed better than any combination of the neural predictor with any other
640 predictors based on self-report data (Fig. 7). Moreover, it performed slightly better than a
641 model including the interaction of 'social cognition' and 'PPI' (BIC = 65, $R^2 = 0.60$, model
642 10 in Fig. 7), which revealed no significant interaction and therefore no moderation effect
643 ($\beta = -0.40$, $t_{(49)} = -0.38$, $p = 0.702$, 95% CI [-2.52, 1.71]). When we added the other self-
644 report predictors (model 13 in Fig. 7), the main effects of 'social cognition' and 'PPI'
645 remained robust, whereas the other predictors showed no significant effect. Accordingly,
646 the increase of connectivity between the right IPL and right dIPFC in the Controlled
647 compared with the Free condition explains variance in individual control-averse behavior
648 that exceeds model predictions based on self-report data.

649

650 *The connectivity between right IPL and dIPFC partially mediates the association of social*
651 *cognition with control-averse behavior*

652 After having identified social cognition and the right IPL-dIPFC connectivity as the best
653 predictors of individual control-averse behavior, we asked whether the connectivity might
654 reflect the mechanism through which these social cognitions affect control-averse
655 behavior and therefore capture joint variance. To test this question, we ran a mediation
656 analysis using a three-variable path model (Fig. 8, Baron and Kenny, 1986; Wager et al.,
657 2008), in which the predictor is 'social cognition', the mediator is the subject-wise beta
658 estimate of the Controlled PPI minus the Free PPI regressor between the right IPL and

659 the right dIPFC, and the outcome is the individual control-averse behavior. Following
660 convention (Baron and Kenny, 1986), we considered the mediation to be significant if
661 three conditions were met: the predictor must be related to the mediator (path a), the
662 mediator must be related to the outcome after controlling for the predictor (path b), and
663 the mediation effect, i.e. the product of the a and b path coefficients ($a*b = c-c'$), must
664 be significant. The mediation analysis revealed that the relationship between social
665 cognition and control-averse behavior is partially mediated by the connectivity between
666 right IPL and right dIPFC, i.e. the mediator significantly reduces the association between
667 predictor and outcome (total effect, path c), but the predictor still explains significant
668 variance of the outcome (direct effect, path c', Fig. 8). In other words, the right IPL-
669 dIPFC connectivity explains a significant part of the relationship between social cognition
670 and control-averse behavior, but the predictor and mediator each also explain
671 independent variance.

672

673

674 **DISCUSSION**

675 People value their freedom of choice highly. Interestingly, though, if another person tries
676 to restrict one's choice, some people will comply, whereas others will act against the
677 restriction. These individual differences in control-averse behavior have been well
678 documented, but their driving factors have remained a puzzle. Previous work has
679 suggested several potential predictors of control-averse behavior based on self-reports.
680 To date, however, we know very little about the mechanisms that underlie control-averse
681 behavior at the neural level. Here, we identify a neural mechanism that complements
682 and exceeds self-reported social cognitions, affects and motivations in explaining
683 individual differences in control-averse behavior.

684 To do so, we combined fMRI with a Control aversion task (Falk and Kosfeld, 2006;
685 Schmelz and Ziegelmeyer, 2015), in which subjects' freedom of choice is controlled by
686 another person, and subjects' subsequent monetary allocation to that person serves as
687 a measure of control-averse behavior. Specifically, we aimed to identify neural
688 mechanisms that could explain individual differences in control-averse behavior. Our
689 results both replicate prior behavioral studies and provide novel insights into the
690 neurobiological basis of control-averse behavior. We replicated that control of one's
691 freedom of choice by another person reduces the willingness to allocate money to that
692 person (Falk and Kosfeld, 2006; Schmelz and Ziegelmeyer, 2015). This effect was
693 augmented in subjects who had little understanding for the other person's behavior or
694 who perceived the restriction of their freedom of choice as a signal of distrust in their
695 intrinsic motivation to choose a generous and fair allocation (Falk and Kosfeld, 2006).
696 We also found that control-averse behavior was accompanied by negative affects
697 (Dillard and Shen, 2005) and the motivation to restore one's freedom of choice (Brehm,
698 1966; Miron and Brehm, 2006). This is in line with previous research on reactance that
699 has focused on behavioral intentions in hypothetical scenarios (Sittenthaler et al., 2015)
700 or behavior in non-social settings (Chartrand et al., 2007). Our study complements and
701 extends this research by providing evidence of the motivation to act against the
702 restriction of one's freedom of choice during social decisions with actual consequences.
703 A direct comparison of the predictors based on the self-report data revealed that a
704 combination of the social cognitions perceived distrust and understanding explained
705 individual control-averse behavior best at the behavioral level.
706 At the neural level, we found that control-averse behavior could be predicted by
707 functional connectivity between the right IPL and the bilateral dlPFC/ middle frontal
708 gyrus. Notably, our finding is specific to the right IPL, which corroborates previous work

709 examining its role in subjective choice restrictions (Filevich et al., 2013). The
710 involvement of both IPL and dlPFC in control-averse behavior could be attributed to their
711 functions suggested in previous neuroimaging studies. The IPL has traditionally been
712 associated with the reorienting of attention to both social and non-social stimuli (Corbetta
713 et al., 2008) as well as number processing (Dehaene et al., 2003). Also, more recent
714 work has linked the IPL to social distance encoding (Chiao et al., 2009; Parkinson et al.,
715 2014), suggesting that the IPL might perform analogous operations in visuospatial and
716 social contexts (Yamazaki et al., 2009; Parkinson et al., 2014). Therefore, it seems
717 plausible that the differential IPL activation during decisions in the Controlled compared
718 with the Free condition might reflect the encoding of or attention reorientation to the
719 context (i.e. being controlled or not) that is relevant for the decision (i.e. to counteract or
720 not). The differential IPL activation alone, however, did not explain individual differences
721 in control-averse behavior, suggesting that the IPL encodes the difference between the
722 Controlled and the Free condition irrespective of the subjects' individual control aversion.
723 Instead, individual differences in control-averse behavior could be explained by the
724 connectivity of right IPL with the dlPFC, two regions that are directly connected through
725 fiber tracts (Mars et al., 2012). Moreover, the IPL and regions in the lateral PFC show
726 robust intrinsic functional coupling (Mars et al., 2011) as well as increased task-based
727 coupling during changes of choice strategy (Daw et al., 2006; Boorman et al., 2009).
728 Follow-up studies could investigate whether individual differences in anatomical or
729 resting state functional connectivity between the IPL and dlPFC might contribute to
730 control-averse behavior.

731 The dlPFC has been commonly associated with cognitive control (MacDonald et al.,
732 2000; Miller and Cohen, 2001) and overcoming conflicts in decisions that require self-
733 control (Knoch et al., 2006; Hare et al., 2009; Figner et al., 2010; Baumgartner et al.,

734 2011). Correspondingly, the notion that control-averse behavior requires cognitive
735 control is supported by our behavioral data: Although all subjects demonstrated an
736 intrinsic motivation to choose a high level, control-averse subjects were more likely to
737 dislike the restriction of their freedom of choice and to feel the urge to use their
738 remaining freedom of choice. This suggests that control-averse subjects perceived the
739 decisions in the Controlled condition as a conflict between the general motivation to
740 choose a high level and the condition-specific motivation to act against the restriction.
741 Given its suggested role in cognitive control, this could explain why the dIPFC was more
742 strongly recruited by control-averse subjects during decisions in the Controlled condition
743 as indicated by the connectivity analysis and illustrated in the time course plots.
744 Furthermore, model comparisons indicate that the right IPL-dIPFC connectivity explains
745 additional variance of the individual control-averse behavior that has remained
746 unexplained by self-reports alone. More specifically, we find that the neural data
747 complement the self-reports of social cognitions. Together, these two predictors explain
748 a sizable amount of variance in the control-averse behavior and provide the best data fit
749 among the tested models. Notably, the IPL cluster that we find lies in close proximity to
750 the temporoparietal junction (Mars et al., 2012; Igelstrom et al., 2015), which is
751 considered a key region in social cognition (Decety and Lamm, 2007; Cabeza et al.,
752 2012; Carter and Huettel, 2013; Krall et al., 2015). It has been proposed that the IPL
753 shares information with the temporoparietal junction via joint connections in the dIPFC/
754 middle frontal gyrus (Corbetta et al., 2008), matching the target region of our connectivity
755 analysis. In line with this notion, we found that the right IPL-dIPFC connectivity partially
756 mediates the association between social cognition and control-averse behavior. The
757 partial mediation and model comparisons further suggest that the right IPL-dIPFC
758 connectivity explains variance that could not be captured by self-reports. This

759 emphasizes once more that for a comprehensive understanding of a complex human
760 behavior such as control-averse behavior, it is essential to incorporate
761 neurophysiological factors. Although the IPL and dIPFC certainly have intricate roles in
762 decision making, together our data provide evidence that the Controlled condition
763 represents a socially salient event and that the right IPL-dIPFC connectivity might
764 contribute to the integration of social cognition into control-averse behavior.
765 Lastly, it is important to acknowledge limitations of our study and provide suggestions on
766 how to address them in future work. First, it would be interesting to see whether our
767 results generalize to non-social scenarios. Falk and Kosfeld (2006) have demonstrated,
768 however, that replacing the player A with a computer algorithm eliminates control-averse
769 behavior, suggesting that the aversion to the choice restriction might be confounded with
770 the social aspect in our task. Therefore, designing a study that analogously varies the
771 degree of choice restrictions in both a social and non-social context could be an
772 interesting future endeavor.

773 Furthermore, we opted for a small number of trials to increase credibility and limit
774 possible habituation and attention biases. This means that, while our neuroimaging
775 results survive whole-brain correction, some brain activation might have gone
776 undetected. Using a greater number of trials, however, would have come at the risk of a
777 less robust measure of control-averse behavior. In the current data, the robustness of
778 our measure of control-averse behavior is supported by the consistent correlations with
779 the affect and self-report ratings. Similar sanity checks should be incorporated in future
780 neuroimaging studies on control-averse behavior.

781 To sum up, this study provides first insights into the neural drivers of individual
782 differences in control-averse behavior, a social phenomenon that is ubiquitous in our
783 society. The prevalence of control-averse behavior and its potential negative

784 consequences have become evident in previous behavioral studies. Advancing our
785 understanding of the mechanisms that give rise to individual differences in control-
786 averse behavior therefore represents an important research goal. Here, we have
787 approached this goal by identifying a neural mechanism that can explain individual
788 differences in control-averse behavior. Our results suggest that a key driver of control-
789 averse behavior is the connectivity between brain regions that are reliably, albeit not
790 exclusively, involved in attention reorientation and cognitive control. This connectivity
791 complements what could be measured by self-reports alone and thereby improves our
792 understanding of the mechanisms underlying control-averse behavior. While more work
793 is needed to investigate the exact neural computations and extend these findings to
794 more complex social interactions, this study has brought us a significant step forward in
795 unraveling the drivers of individual differences in control-averse behavior.
796
797

798 **REFERENCES**

- 799 Baron RM, Kenny DA (1986) The Moderator-Mediator Variable Distinction in Social
800 Psychological Research: Conceptual, Strategic, and Statistical Considerations. *J*
801 *Pers Soc Psychol* 51:1173–1182.
- 802 Baumgartner T, Knoch D, Hotz P, Eisenegger C, Fehr E (2011) Dorsolateral and
803 ventromedial prefrontal cortex orchestrate normative choice. *Nat Neurosci*
804 14:1468–1474.
- 805 Boorman ED, Behrens TEJ, Woolrich MW, Rushworth MFS (2009) How green is the
806 grass on the other side? Frontopolar cortex and the evidence in favor of alternative
807 courses of action. *Neuron* 62:733–743.
- 808 Bradley M, Lang PJ (1994) Measuring emotion: the Self-Assessment Manikin and the
809 semantic differential. *J Behav Ther Exp Psychiatry* 25:49–59.
- 810 Brehm JW (1966) A theory of psychological reactance. New York: Academic Press.
- 811 Brett M, Anton J-L, Valabregue R, Poline J-B (2002) Region of interest analysis using an
812 SPM toolbox. *Neuroimage* 16:CD-ROM.
- 813 Cabeza R, Ciaramelli E, Moscovitch M (2012) Cognitive contributions of the ventral
814 parietal cortex: An integrative theoretical account. *Trends Cogn Sci* 16:338–352.
- 815 Carter RM, Huettel SA (2013) A nexus model of the temporal-parietal junction. *Trends*
816 *Cogn Sci* 17:328–336.
- 817 Chartrand TL, Dalton AN, Fitzsimons GJ (2007) Nonconscious relationship reactance:
818 When significant others prime opposing goals. *J Exp Soc Psychol* 43:719–726.
- 819 Chiao JY, Harada T, Oby ER, Li Z, Parrish T, Bridge DJ (2009) Neural representations
820 of social status hierarchy in human inferior parietal cortex. *Neuropsychologia*
821 47:354–363.
- 822 Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain:

- 823 from environment to theory of mind. *Neuron* 58:306–324.
- 824 Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for
825 exploratory decisions in humans. *Nature* 441:876–879.
- 826 De las Cuevas C, Peñate W, Betancort M, de Rivera L (2014) Psychological reactance
827 in psychiatric patients: Examining the dimensionality and correlates of the Hong
828 Psychological Reactance Scale in a large clinical sample. *Pers Individ Dif* 70:85–
829 91.
- 830 Decety J, Lamm C (2007) The role of the right temporoparietal junction in social
831 interaction: How low-level computational processes contribute to meta-cognition.
832 *Neuroscientist* 13:580–593.
- 833 Dehaene S, Piazza M, Pinel P, Cohen L (2003) Three parietal circuits for number
834 processing. *Cogn Neuropsychol* 20:487–506.
- 835 Dillard JP, Shen L (2005) On the nature of reactance and its role in persuasive health
836 communication. *Commun Monogr* 72:144–168.
- 837 Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM (1993) 3D statistical
838 neuroanatomical models from 305 MRI volumes. In: *IEEE Conference Record
839 Nuclear Science Symposium and Medical Imaging Conference*, pp 1813–1817.
840 IEEE.
- 841 Falk A, Kosfeld M (2006) The hidden costs of control. *Am Econ Rev* 96:1611–1630.
- 842 Figner B, Knoch D, Johnson EJ, Krosch AR, Lisanby SH, Fehr E, Weber EU (2010)
843 Lateral prefrontal cortex and self-control in intertemporal choice. *Nat Neurosci*
844 13:538–539.
- 845 Filevich E, Vanneste P, Brass M, Fias W, Haggard P, Kühn S (2013) Brain correlates of
846 subjective freedom of choice. *Conscious Cogn* 22:1271–1284.
- 847 Gläscher J (2009) Visualization of group inference data in functional neuroimaging.

- 848 Neuroinformatics 7:73–82.
- 849 Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves
850 modulation of the vmPFC valuation system. *Science* 324:646–648.
- 851 Igelstrom KM, Webb TW, Graziano MSA (2015) Neural Processes in the Human
852 Temporoparietal Cortex Separated by Localized Independent Component Analysis.
853 *J Neurosci* 35:9432–9445.
- 854 Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E (2006) Diminishing reciprocal
855 fairness by disrupting the right prefrontal cortex. *Science* 314:829–832.
- 856 Krall SC, Rottschy C, Oberwelland E, Bzdok D, Fox PT, Eickhoff SB, Fink GR, Konrad K
857 (2015) The role of the right temporoparietal junction in attention and social
858 interaction as revealed by ALE meta-analysis. *Brain Struct Funct* 220:587–604.
- 859 MacDonald AW, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the
860 dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*
861 288:1835–1838.
- 862 Mars RB, Jbabdi S, Sallet J, O'Reilly JX, Croxson PL, Olivier E, Noonan MP, Bergmann
863 C, Mitchell AS, Baxter MG, Behrens TEJ, Johansen-Berg H, Tomassini V, Miller KL,
864 Rushworth MFS (2011) Diffusion-Weighted Imaging Tractography-Based
865 Parcellation of the Human Parietal Cortex and Comparison with Human and
866 Macaque Resting-State Functional Connectivity. *J Neurosci* 31:4087–4100.
- 867 Mars RB, Sallet J, Schüffelgen U, Jbabdi S, Toni I, Rushworth MFS (2012) Connectivity-
868 based subdivisions of the human right “temporoparietal junction area”: Evidence for
869 different areas participating in different cortical networks. *Cereb Cortex* 22:1894–
870 1903.
- 871 Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev*
872 *Neurosci* 24:167–202.

- 873 Miron AM, Brehm JW (2006) Reactance Theory-40 years later. *Zeitschrift Für*
874 *Sozialpsychologie* 37:9–18.
- 875 Parkinson C, Liu S, Wheatley T (2014) A Common Cortical Metric for Spatial, Temporal,
876 and Social Distance. *J Neurosci* 34:1979–1987.
- 877 Rudolf S, Hare TA (2014) Interactions between dorsolateral and ventromedial prefrontal
878 cortex underlie context-dependent stimulus valuation in goal-directed choice. *J*
879 *Neurosci* 34:15988–15996.
- 880 Schmelz K, Ziegelmeyer A (2015) Social distance and control aversion: Evidence from
881 the internet and the laboratory. *Thurgau Inst Econ Res Pap Ser* 100:1–25.
- 882 Schwenkmezger P, Hodapp V, Spielberger C (1992) *Das State-Trait-Ärgerausdrucks-*
883 *Inventar (STAXI)*. Bern: Huber.
- 884 Sittenthaler S, Traut-Mattausch E, Steindl C, Jonas E (2015) Salzburger State
885 Reactance Scale (SSR Scale): Validation of a scale measuring state reactance. *Z*
886 *Psychol* 223:257–266.
- 887 Spielberger CD (1988) *State-Trait Anger Expression Inventory (STAXI)*. Odessa, FL:
888 Psychological Assessment Resources.
- 889 Van Petegem S, Soenens B, Vansteenkiste M, Beyers W (2015) Rebels with a cause?
890 Adolescent defiance from the perspective of reactance theory and self-
891 determination theory. *Child Dev* 86:903–918.
- 892 Wager TD, Davidson ML, Hughes BL, Lindquist M a, Ochsner KN (2008) Prefrontal-
893 subcortical pathways mediating successful emotion regulation. *Neuron* 59:1037–
894 1050.
- 895 Yamazaki Y, Hashimoto T, Iriki A (2009) The posterior parietal cortex and non-spatial
896 cognition. *F1000 Biol Rep* 6:1:74.
- 897 Ziegelmeyer A, Schmelz K, Ploner M (2012) Hidden costs of control: Four repetitions

898 and an extension. Exp Econ 15:323–340.

899

900 **FIGURE LEGENDS**

901

902 **Figure 1. Control aversion task.** For every trial, the subject is presented with the
903 decision from a new player A and the available generosity levels. Each generosity level
904 represents an allocation of monetary units between the player A (top value) and the
905 subject (bottom value). In the Free condition (blue frame), the player A lets the subject
906 choose freely between level one to five (from left to right). In the Controlled condition
907 (orange frame), the player A requests a minimum of level two and thereby restricts the
908 subject's choice to the levels two to five. The decision window that is highlighted in the
909 figure is defined as the time between the onset of the choice options and the initial
910 movement of the red selection frame. Lastly, the subject is presented with three pictorial
911 assessment scales, which range from unhappy to happy (left to right), from calm to
912 angry, and from being controlled to having control. The durations of the fixation displays
913 were jittered.

914

915 **Figure 2. Choice behavior. A-B**, boxplots of the ratings of having control and
916 chosen generosity levels, respectively, in the Controlled and the Free condition.
917 The central mark of each box shows the median, the box edges show the 25th and
918 75th percentiles, and the whiskers represent the limit beyond which a data point is
919 considered an outlier (denoted as cross). The connected data points in the center
920 show individual subject's means. **C**, The histograms show the distribution of
921 subjects' mean and variance of chosen levels in the Controlled and the Free
922 condition. Data from $n = 51$ subjects are shown.
923
924

925 **Figure 3. Correlation of control-averse behavior with negative affects,**
926 **perceived distrust, understanding, freedom restoration and fairness. A,** mean
927 unhappiness and anger ratings in the Controlled minus the Free condition and
928 individual tendencies for outward directed anger expression, respectively, plotted
929 against the individual control-averse behavior, computed as the difference between
930 the mean chosen level in the Free minus the Controlled condition. **B,** individual
931 ratings of perceived distrust, understanding, freedom restoration and fairness
932 plotted against individual control-averse behavior. Observations are jittered along
933 the x-axis to reduce overlap for visualization. Regression lines were fitted with
934 bisquare robust regressions. Data from n = 51 subjects are shown.
935

936 **Figure 4. Models based on self-report data.** These diagrams show seven models
937 predicting individual control-averse behavior (y), based on self-reports of social
938 cognition (S), freedom restoration (F), and negative affect (A). Arrows indicate main
939 effects. The bar graphs show the Bayesian Information Criterion (BIC) and R^2 for
940 each model, with the winning model highlighted in black.
941

942 **Figure 5. Connectivity between right IPL and dIPFC predicts individual**
943 **differences in control-averse behavior.** The figure illustrates that the functional
944 connectivity during decisions in the Controlled as opposed to the Free condition
945 (Controlled PPI – Free PPI) between the right IPL (seed) and regions in the dIPFC/
946 middle frontal gyrus and the posterior parietal cortex increases as a function of
947 individual control-averse behavior. Left, Statistical parametric maps of the covariate
948 analysis are shown, color coded for the t values as indicated by the color bar,
949 thresholded at $p_{FWE} < 0.05$, and projected on a template brain in MNI space. Right,
950 Graphs show the individual level of control-averse behavior (x-axes) plotted against
951 the single-subject means of the beta estimates extracted from the functional
952 clusters in the right and left dIPFC (circled on the left) for the Controlled PPI – Free
953 PPI effect, the Controlled PPI effect and the Free PPI effect seeded in the right IPL
954 (y-axes). Observations are jittered along the x-axis to reduce overlap for
955 visualization. Regression lines were fitted with bisquare robust regressions. Data
956 from $n = 51$ subjects are shown. R, right; L, left; IPL, inferior parietal lobule; dIPFC,
957 dorsolateral prefrontal cortex.
958

959 **Figure 6. BOLD time courses of decisions in the Controlled and Free**
960 **condition.** The IPL shows a similar pattern for not control-averse subjects (with
961 levels of control-averse behavior ≤ 0 , $n = 10$) and control-averse subjects (with
962 levels of control-averse behavior > 0 , $n = 41$), whereas the dIPFC shows a distinct
963 pattern for control-averse subjects. The graphs show averaged time courses of
964 BOLD activation in the bilateral IPL (top row) and the bilateral dIPFC/ middle frontal
965 gyrus (bottom row) for decisions in the Controlled (orange) and the Free condition
966 (blue). The brain maps in the center depict the search volumes used for the time
967 course extractions. The horizontal lines at the top of the graphs indicate time points
968 at which the conditions differ significantly (Wilcoxon signed rank test, two-tailed,
969 $p < 0.05$). The dashed vertical lines mark the onset of the decision window, at which
970 the time courses were mean-corrected. The transparent areas show standard
971 errors of the mean. Note that these plots were not used to infer the main effect of
972 Controlled $>$ Free condition. L, left; R, right; IPL, inferior parietal lobule; dIPFC,
973 dorsolateral prefrontal cortex.

974 **Figure 7. Models based on self-report and neural data.** These diagrams show
975 seven models predicting individual control-averse behavior (y), based on self-
976 reports of social cognition (S), freedom restoration (F), negative affect (A), and
977 subject-wise estimates of right IPL-dIPFC connectivity in the Controlled minus the
978 Free condition (PPI). Arrows indicate main effects, and the line with a circular
979 endpoint in model 10 indicates an interaction effect. The bar graphs show the
980 Bayesian Information Criterion (BIC) and R^2 for each model, with the winning model
981 highlighted in black.
982

983 **Figure 8. Results of the mediation analysis testing the relationship between**
984 **social cognition, right IPL-dIPFC connectivity and control-averse behavior.**
985 Left, the path model shows the path coefficients with standard errors of the mean
986 in parentheses, significant at * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$. Right,
987 histogram of the bootstrapped distribution of the mediation effect ($a*b = c-c'$). The
988 lighter grey portion of each histogram denotes the 95% confidence interval for
989 the effect. Data from $n = 51$ subjects were included in this analysis.
990

991 **TABLES**

992

993 **Table 1. Regions in which the connectivity for decisions in the Controlled minus**
 994 **the Free condition (Controlled PPI – Free PPI) seeded in the right IPL is positively**
 995 **associated with individual control-averse behavior.**

Region	Side	MNI coordinates			Cluster size k_E	Max stat t	p_{FWE}
		x	y	z			
dIPFC/ middle frontal gyrus	R	42	47	22	105	4.88	<0.001
		24	50	5		4.58	
		48	35	29		4.16	
Angular gyrus	L	-33	-55	36	411	4.80	<0.001
		6	-70	50		4.67	
		27	-73	50		4.40	
Precuneus	R	18	-67	29	40	4.80	0.047
		3	-67	29		3.10	
		21	-58	26		3.01	
dIPFC/ middle frontal gyrus	L	-45	29	29	41	4.54	0.042
		-39	38	26		3.10	
		-45	35	19		2.95	
IPL	L	-39	-52	57	43	4.22	0.033
		-33	-58	57		3.35	
		-24	-64	60		2.95	

996 Results from the covariate analysis are shown (sample size $n = 51$ subjects). Height
 997 threshold $t_{(49)} = 2.68$, extent threshold $k_E > 40$. All activations survive whole-brain

998 correction for multiple comparisons based on FWE-control at the cluster level. dIPFC,
 999 dorsolateral prefrontal cortex; IPL, inferior parietal lobule.

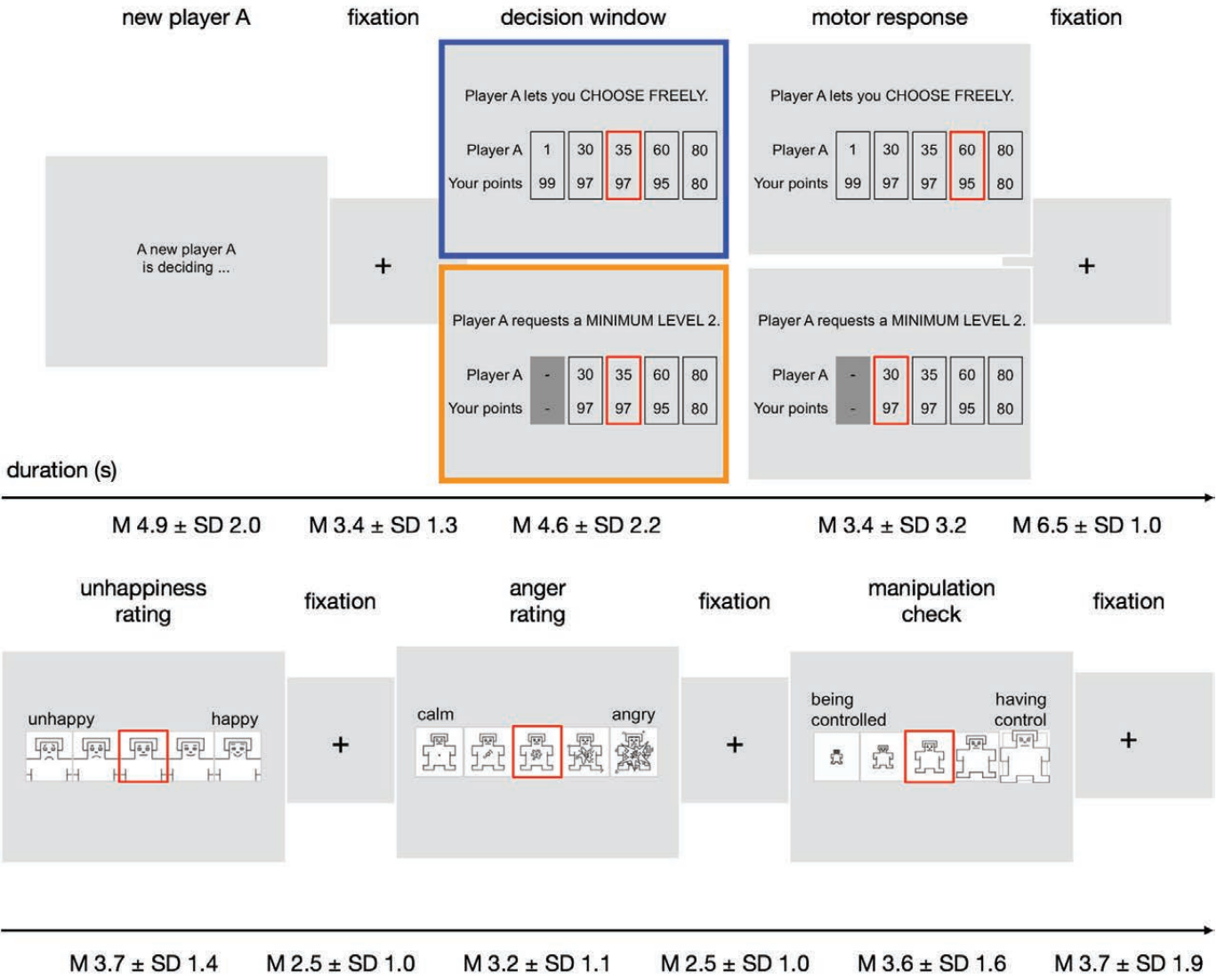
1000

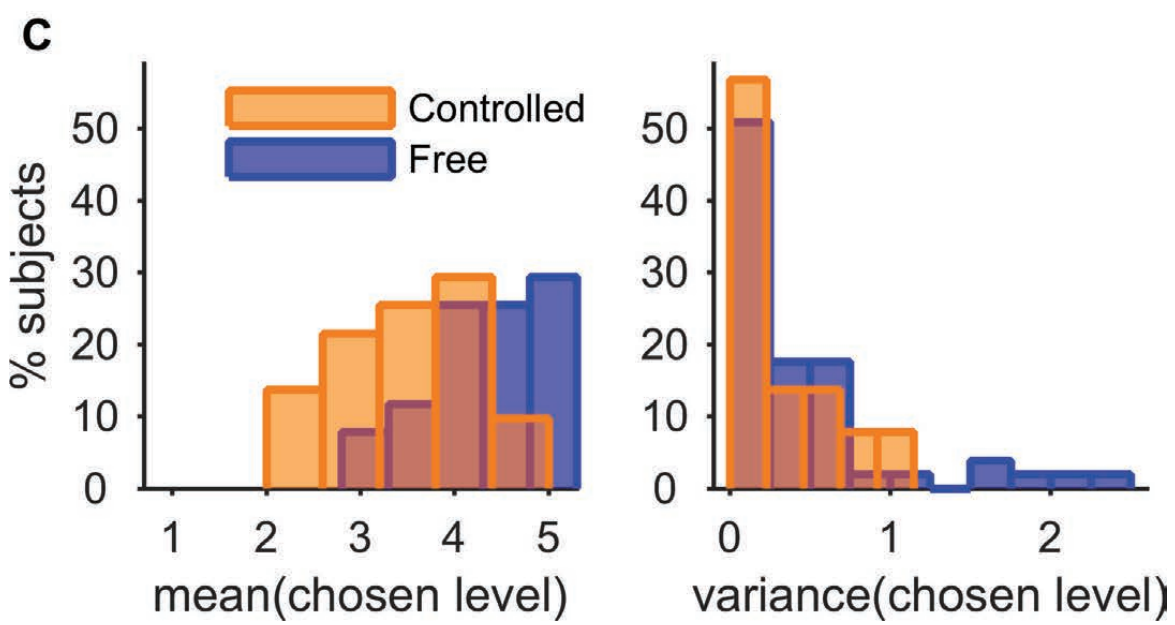
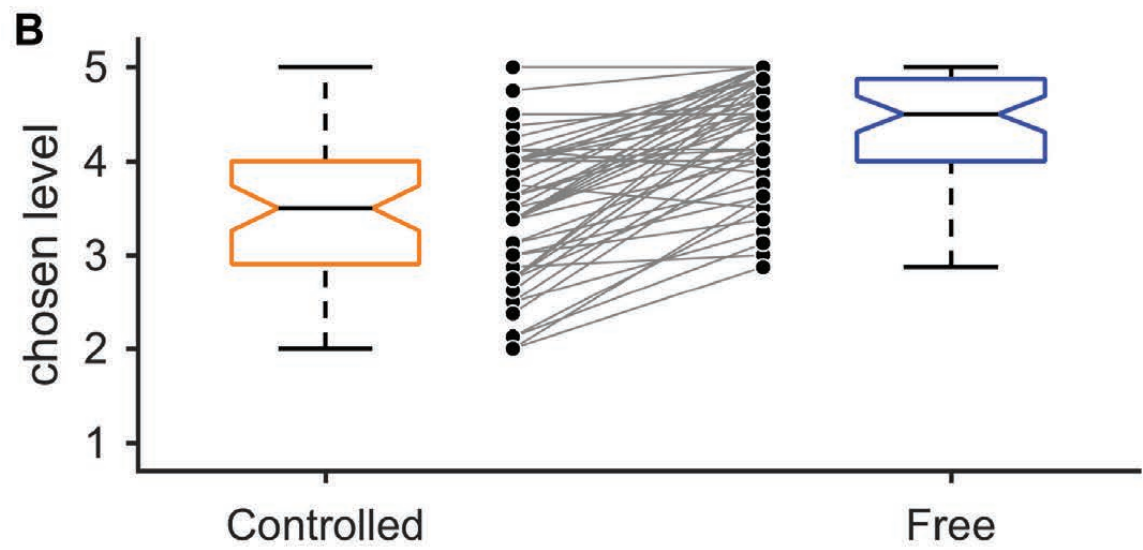
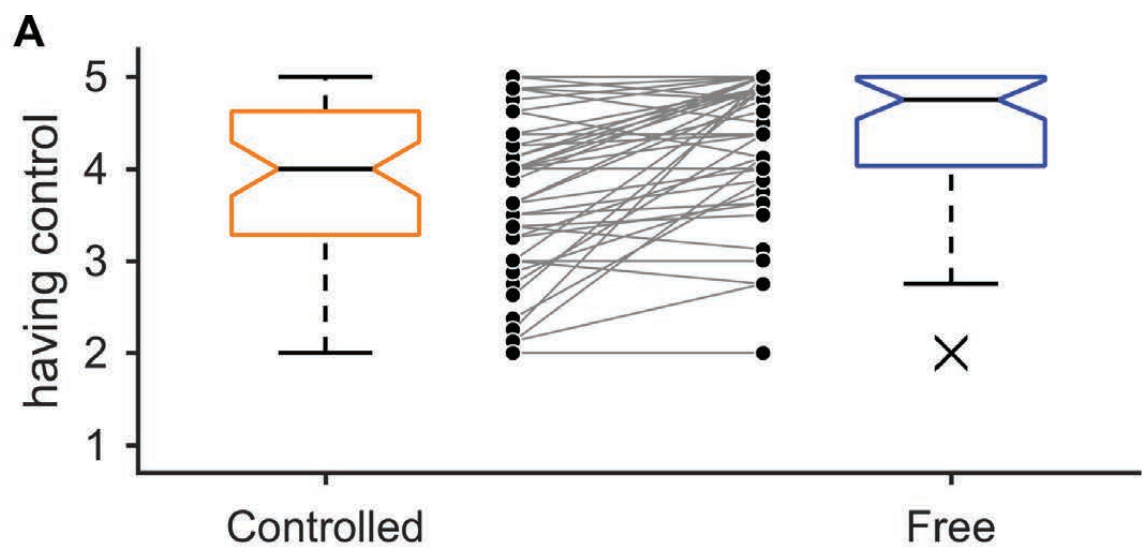
1001

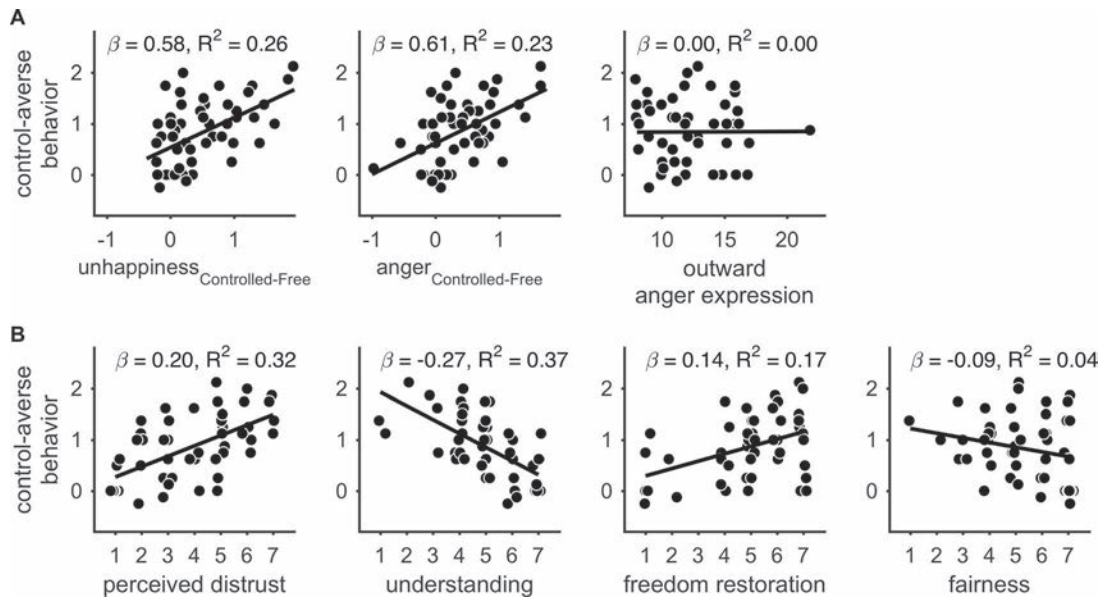
1002 **Table 2. Model comparison. Individual differences in control-averse behavior**
 1003 **predicted by social cognition and right IPL-dIPFC connectivity in the Controlled**
 1004 **minus the Free condition (models 1 and 9 in Fig. 7).**

Dependent variable: control-averse behavior												
model 1							model 9					
					95% CI						95% CI	
	β	SE	t	p	Lower	Upper	β	SE	t	p	Lower	Upper
Cognition	1.36	0.19	7.19	< 0.001	0.98	1.74	1.06	0.20	5.39	< 0.001	0.66	1.45
IPL-dIPFC connectivity							0.92	0.28	3.28	0.002	0.36	1.49
(Intercept)	0.84	0.06	13.51	< 0.001	0.72	0.97	0.43	0.14	3.16	0.003	0.16	0.71
BIC	68.1						61.7					
R ²	0.51						0.60					
Observations	51											

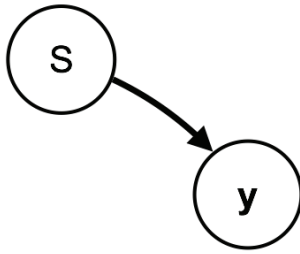
1005



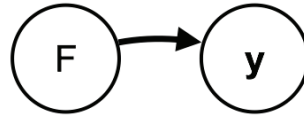




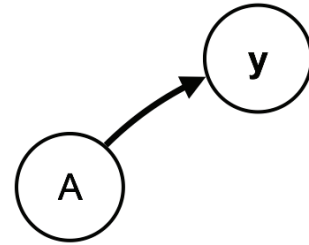
model 1



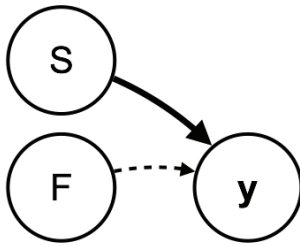
model 2



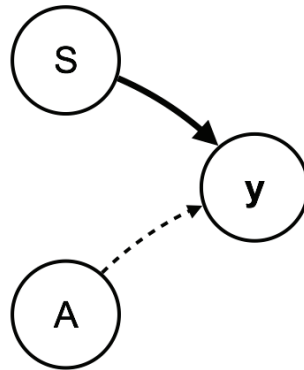
model 3



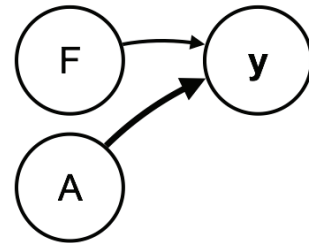
model 4



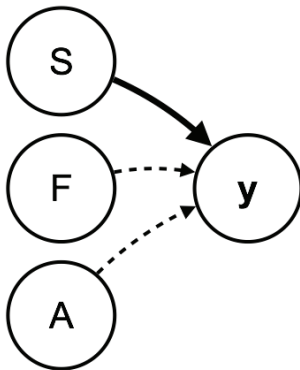
model 5



model 6

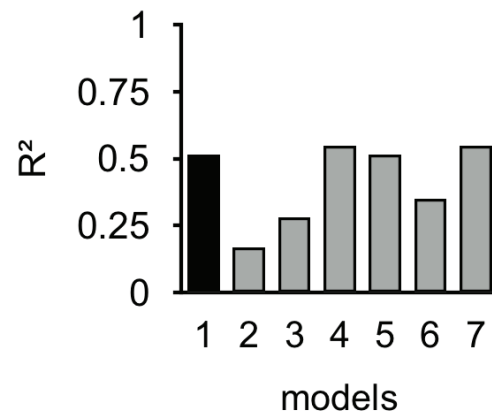
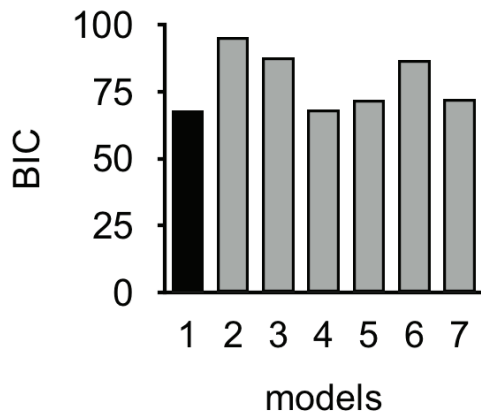


model 7

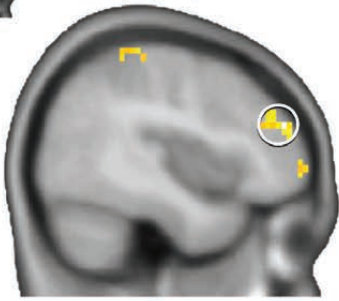


— $p < 0.005$
 — $p < 0.05$
 - - - $p > 0.05$

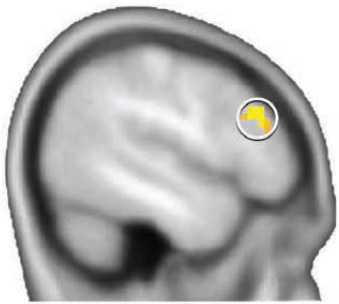
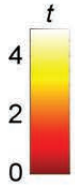
y = control-averse behavior
 S = social cognition
 F = freedom restoration
 A = negative affect



seed R IPL

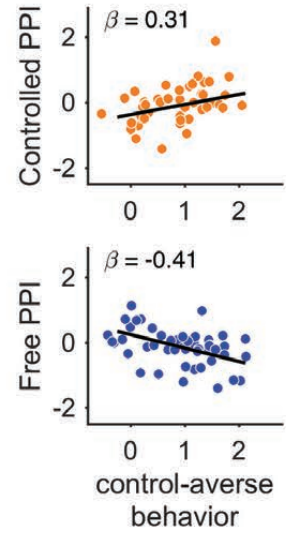
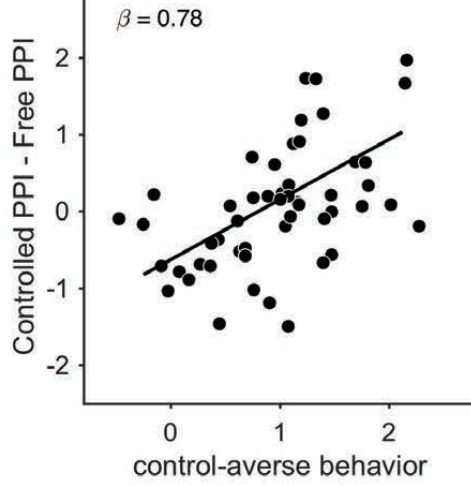


x = 42

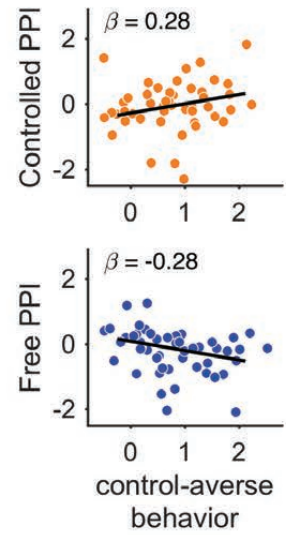
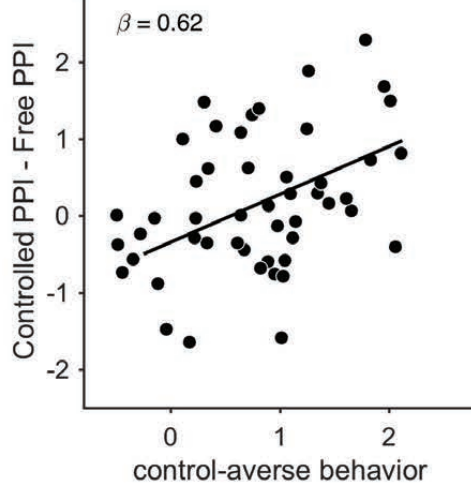


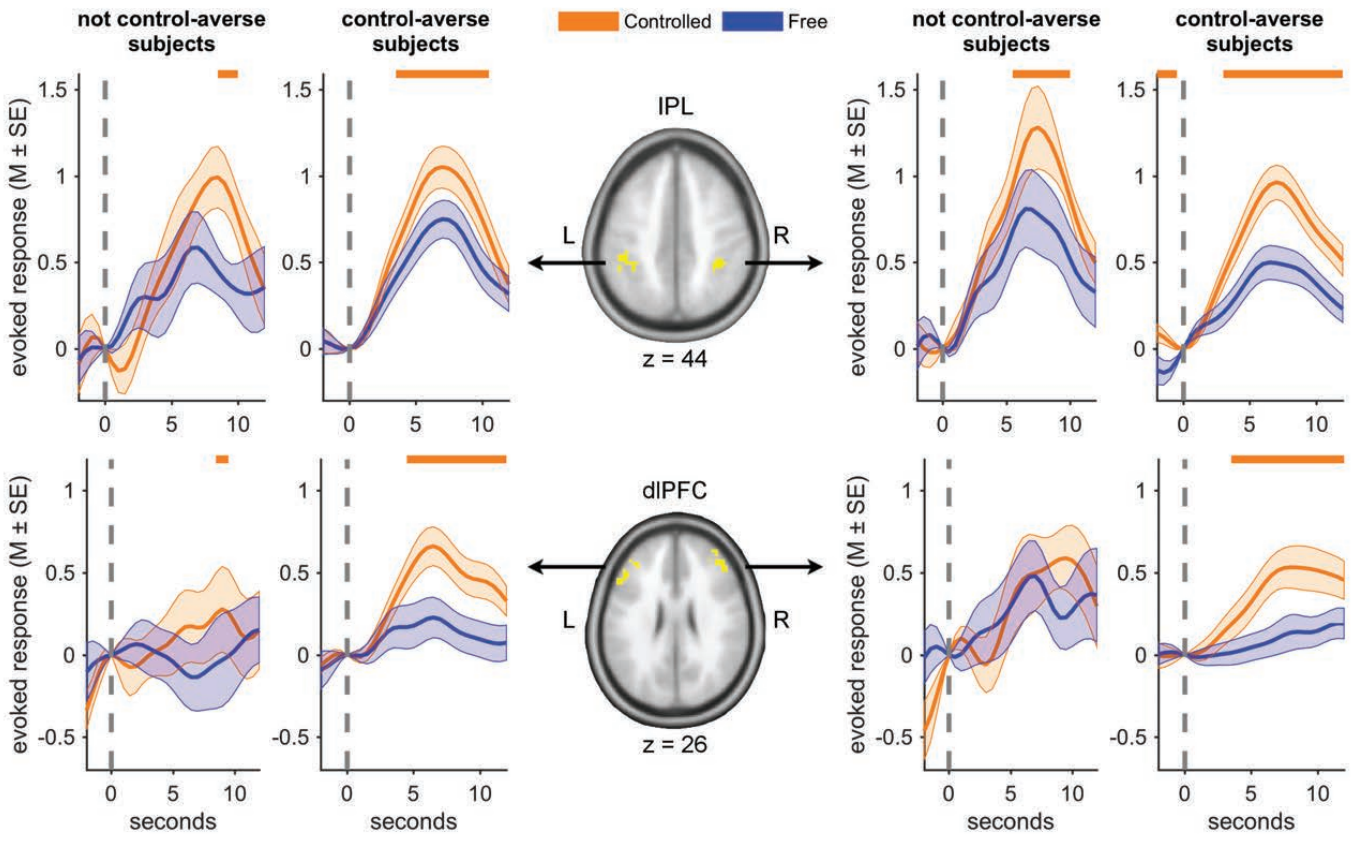
x = -48

R IPL - R dIPFC coupling

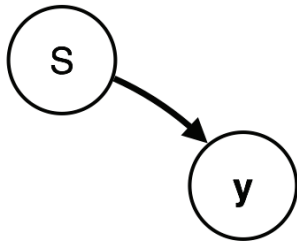


R IPL - L dIPFC coupling

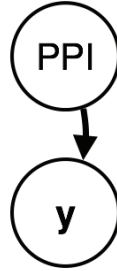




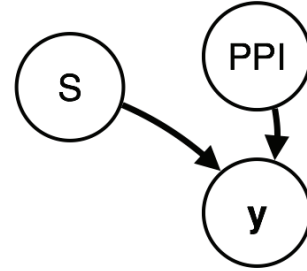
model 1



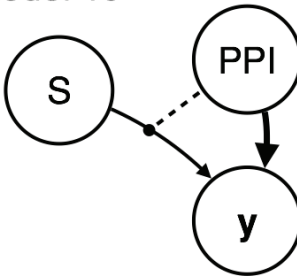
model 8



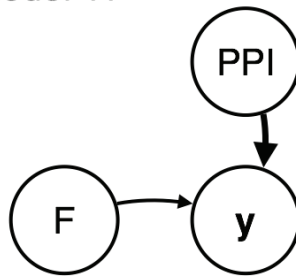
model 9



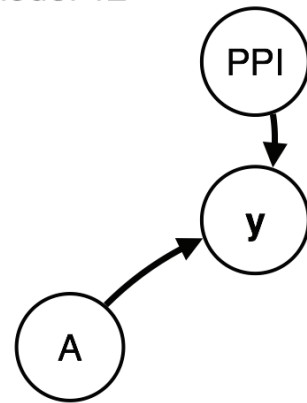
model 10



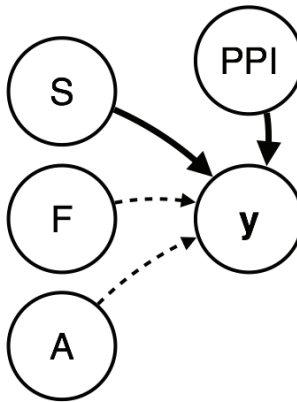
model 11



model 12



model 13



— $p < 0.005$
 — $p < 0.05$
 - - - $p > 0.05$

y = control-averse behavior
 S = social cognition
 F = freedom restoration
 A = negative affect
 PPI = IPL-dIPFC connectivity

