Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/5/3/eaau3413/DC1

Supplementary Materials for

The neural circuitry of affect-induced distortions of trust

Jan B. Engelmann, Friederike Meyer, Christian C. Ruff, Ernst Fehr*

*Corresponding author. Email: Ernst.fehr@econ.uzh.ch (E.F.); Jbengelmann@gmail.com (J.B.E.)

Published 13 March 2019, *Sci. Adv.* **5**, eaau3413 (2019)
DOI: 10.1126/sciadv.aau3413

This PDF file includes:

1 Supplementary Analyses

Section S1. Ordinary least squares regression investigating the influence of experienced electrical stimulation on choice

It is possible that the main driving force of the behavioral change that we report in the main paper is not due to the affective impact of our threat of shock manipulation, but to the actual experience of electrical stimulation immediately prior to decision-making. To assess this possibility, we investigated the effects of experienced electrical stimulation on choice behavior in the trust and the non-social control game. To this end, we ran comprehensive ordinary least-squares (OLS) regression analyses in R version 2.5.12. These analyses predicted for each individual *i* the observed choice $T_{i,t}$ on trial *t* with the following equation:

$$T_{i,t} = \beta_0 + \beta_1 * \text{Threat}_{i,t} + \beta_2 * \text{Game type}_{i,t} + \beta_3 * \text{US}_{i,t} + \beta_4 * \text{PS}_{i,t} + \beta_5 * \text{Threat}_{i,t} * \text{US}_{i,t} + \beta_6 * \text{Game type}_{i,t} \text{ x } \text{US}_{i,t} + \beta_7 * \text{Threat}_{i,t} \text{ x Game type}_{i,t} \text{ x } \text{US}_{i,t} + \beta_8 * \text{Threat}_{i,t} \text{ x } \text{PS}_{i,t} + \beta_9 * \text{Game type}_{i,t} \text{ x } \text{PS}_{i,t} + \beta_{10} * \text{Threat}_{i,t} \text{ x Game type}_{i,t} \text{ x } \text{PS}_{i,t} + \beta_{11} * X_i + \varepsilon_{i,t} \qquad \text{(eq. 1)}$$

$T_{i,t}$ reflects the transfer amounts to either a lottery or Player B on a given trial. *Threat*$_{i,t}$ is an effect-coded variable (1 = threat of shock, -1 = safety) reflecting the *expectation* of impending painful electric shock during the current choice scenario, not the *experience* of shock. *Game type*$_{i,t}$ is an effect-coded variable (1 = trust game, -1 = control game) reflecting transfers within the context of a trust game. $US_{i,t}$ and $PS_{i,t}$ are effect-coded variables (1 = shock, -1 = no shock) encoding the *experience* of at least one unpredictable (US) and, respectively, predictable shock (PS) in the interval from 5 seconds prior to the display of the choice scenario until button press. We included PS in our regression models, in order to control for the influence of reminder shocks presented during the block cue. Of note, we also investigated shorter and longer intervals from 1 to 10 seconds prior to the choice scenario in 1-second increments and in all cases only *Threat* and *Game type* reach statistical significance, no other factors or interactions. The model also contains a constant parameter (β_0), which measures the average transfer to the lottery in the no threat condition. Finally, the model includes a set of mean-centered socio-economic control variables (X_i, e.g., age and gender) and relevant interaction terms that reflect differential effects of experiencing shocks on transfer behavior in different Game type and emotion contexts. We employed a random-effects model with robust standard errors adjusted for clustering on the subject level.

The results show significant treatment (threat) and game type (trust) effects, which we also report in the ANOVA analyses in the main paper. At the same time, regression results do not show significant effects of both unpredictable ($p = 0.23$) and predictable ($p = 0.62$) *experienced* electrical stimulation on choice behavior. Specifically, we do not observe a significant effect of, or interaction with, both unpredictable and predictable electrical stimulation. Together, regression results support and extend the results reported in the main paper. Specifically, threat of shock remains a significant predictor of choice behavior ($p < 0.001$) in the trust and the non-social control game, even when controlling for the presence of actually experienced electrical stimulation. Furthermore, we show that the experience of shock does not significantly impact behavior in any of our analyses, indicating that the *expectancy* of shock, not the *experience* of shock, significantly impacted choice behavior.

Table S1. Ordinary least squares regression results reflecting the influence of experienced electrical stimulation on choice behavior.

Dependent Variable: Amount Transferred	(1)		(2)		(3)	
Threat	-0.659	****	-0.5482	****	-0.627	****
	(0.1711)		(0.1389)		(0.1596)	
Game Type	1.2479	****	1.2468	****	1.338	****
	(0.3242)		(0.3094)		(0.3097)	
Game Type * Threat	-0.0128		0.0071		-0.0654	
	(0.1334)		(0.1107)		(0.1185)	
Unpredictable S	0.0146		0.0161			
	(0.1179)		(0.1172)			
Threat * Unpredictable S	-0.0649		-0.0847			
	(0.1254)		(0.1232)			
Game Type * Unpredictable S	-0.1666		-0.167			
	(0.1170)		(0.1186)			
Game Type * Threat * Unpredictable S	0.1002		0.099			
	(0.1413)		(0.1422)			
Predictable S	0.0011				0.0064	
	(0.1237)				(0.1255)	
Threat * Predictable S	-0.1837				-0.1941	
	(0.1099)				(0.1069)	
Game Type * Predictable S	0.0017				-0.0198	
	(0.1035)				(0.1061)	
Game Type * Threat * Predictable S	-0.0303				-0.0159	
	(0.1023)				(0.1032)	
Gender	1.2645		1.2634		1.2659	
	(1.3527)		(1.3533)		(1.3536)	
Age	0.264		0.264		0.2638	
	(0.3253)		(0.3254)		(0.3258)	
Intercept	14.0029	****	14.0041	****	13.9996	****
	(0.7103)		(0.7084)		(0.7103)	
F	17.26		24.78		24.72	
P	<0.001		<0.001		<0.001	
N	3427		3427		3427	
AIC (corrected)	22718.85		22712.35		22712.88	

This table reports OLS coefficient estimates (robust standard errors corrected for clustering on the individual level in parentheses). The dependent variable in both columns is the amount invested in either Player B or the lottery. Column (1) contains the results from the full regression that includes both predictable shocks (i.e., reminder shocks) and unpredictable shocks and all interactions, while columns (2) and (3) reflect results from reduced models that exclude the influence of predictable (2) and unpredictable (3) shocks. We used effect coding for the following dummy variables: (1) “Threat” reflects the *expectation* of impending painful electric shock during the current choice scenario; (2) “Game type” indicates trust compared to control trials (3) “Unpredictable S” and (4) “Predictable S” are dummy variables encoding the experience of at least one unpredictable and, respectively, predictable shock in the interval from 5 seconds prior to the display of the choice screen until button press. We also investigated shorter and longer intervals from 1 to 10 seconds prior to the choice scenario in 1-second steps and in all cases only Threat and Game type reach statistical significance. “Age” is the mean-centered individual’s age in years, and “Gender” is a gender dummy encoding the influence of being male on transfer rates. The number of observations is less than the total number of choice scenarios (3444) due to omissions (response times > 5.5 seconds). Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, **** $p < 0.001$

Section S2. Behavioral differences between trust and control decisions

Despite the similar effects of aversive affect on transfer rates and reaction times across the two game types, three sources of evidence underline that subjects treated decisions in the trust and non-social control games differently and, therefore, that aversive affect impacted separable underlying cognitive processes in the trust and non-social control game. First, average transfer rates were significantly higher in the trust (15.37 (0.767) MU) compared to the non-social control (12.671 (0.804) MU) game as indicated by the significant main effect of game type [$F(1,40) = 20.319$, $p < 0.001$, $\eta^2 = 0.337$]. Moreover, average reaction times were significantly faster in the trust (2.575 (0.07) sec.) compared to the non-social control game (2.675 (0.066) sec.) as indicated by the significant main effect of game type [$F(1,40) = 6.226$, $p = 0.017$, $\eta^2 = 0.135$]. Second, transfer rates within each game type were highly similar across threat conditions, but were significantly different across the two game types within each threat condition. That is, correlation analyses investigating relationships between the transfer rates within and across each game type showed stronger correlations between transfer rates for threat and no threat within each game type (Trust (threat, no threat): $r = 0.931$; NS control (threat, no threat): $r = 0.937$) and somewhat weaker correlations between transfer rates for trust and control decisions within each threat condition (threat (trust, risk): $r = 0.739$; No threat (trust, risk): $r = 0.636$). We compared correlations within (modified Pearson-Filon test, ZPF) and across (Williams t-test) game types using methods outlined in Weaver & Wuensch (2013). This revealed that the correlations between transfer rates in the absence compared to the presence of threat within each game type were similar ($ZPF = -0.23$, $p = 0.816$). However, the correlation between transfer rates in the trust game during threat and no threat was significantly different from (1) the correlation between transfer rates under threat for trust and NS control games ($t(38) = 4.23$, $p < 0.001$) and from (2) the correlation between transfer rates under no threat for trust and NS control games ($t(38) = 5.53$, $p < 0.001$). Third, emotion-related changes in mean transfer rates for the trust game (no threat - threat) did not correlate significantly with emotion-related changes in transfer in the risk game (Pearson’s $r = 0.231$, $p = 0.146$). Together, these results indicate that subjects engaged in different cognitive processes to make decisions in the two game types, supporting the notion that the emotion-induced changes in the two game types reflect the influences of emotions on separable choice-related processes.

Section S3. Assessing complexity differences across game types using choice latency

Using choice latency as a proxy for decision complexity (Krajbich et al. 2010), we tested whether decisions in the trust game were more complex because no information was provided about the repayment probability in the trust game, while a probability range was provided in the control game. At the behavioral level, we inspected choice latency differences between the game types and indeed, we found a small (3.7% of the mean RT) but significant effect of game type on reaction time [F(1,40) = 6.23, p = 0.017), indicating that decisions in the trust game were on average 100ms faster (mean RT in trust game = 2575ms, mean RT in control game = 2675ms). Following the logic that slower RTs reflect more complex decisions (Krajbich et al. 2010), this result indicates that decisions in the trust game are unlikely to have been more complex/opaque compared to decisions in the control game.

We also tested whether our activation patterns overlap with brain regions that are commonly modulated by choice complexity (i.e., reaction times) in similar tasks (Yarkoni et al., 2009). To this end, we conducted a conjunction analysis of our main results and the neurosynth meta-analysis for the term “reaction time” (N = 247 studies). The conjunction analysis does not reveal any overlap with the regions that we observe in the present study (main effect of game type, main effect of task, interaction, all thresholded at $p < 0.001$), including the temporoparietal junction. Thus, our analyses do not support the interpretation that basic differences in choice complexity (as measured by RT) may have caused the activations we observe in the paper.

Section S4. Assessing the lateralization of neuroimaging results

To identify to what extent the effects of aversive affect on the neural correlates of trust decisions are lateralized, we also inspected the main and interaction effects in right TPJ at uncorrected thresholds. This revealed that the right TPJ in fact also showed a numeric (but non-significant) main effect of trust (57, -55, 24, k = 45, p = 0.002) and a numeric, non-significant interaction effect in a superior (62, -54, 28, k = 11, p = 0.004) and an inferior (64, -45, 4, k = 15, p = 0.002) location in pSTS that is part of our TPJ mask. Thus, the effects we observed were somewhat more pronounced in the left hemisphere but not clearly lateralized.

To identify to what extent the functional connectivity effects between TPJ and amygdala were lateralized, we repeated the tests of main and interaction effects in right amygdala at uncorrected thresholds. We found a numeric but non-significant threat-induced disruption of TPJ-amygdala connectivity in the right amygdala (Main effect of threat: 21, -6, -12, k = 5, SV FWE-corrected p = 0.079). Furthermore, we found a numeric, non-significant interaction effect indicating trust-related disruption of TPJ-right amygdala connectivity (30, -7, -12, k = 2, p = 0.008). As for the TPJ, amygdala involvement was more pronounced for the left hemisphere but not clearly lateralized.

2 Supplementary Figures

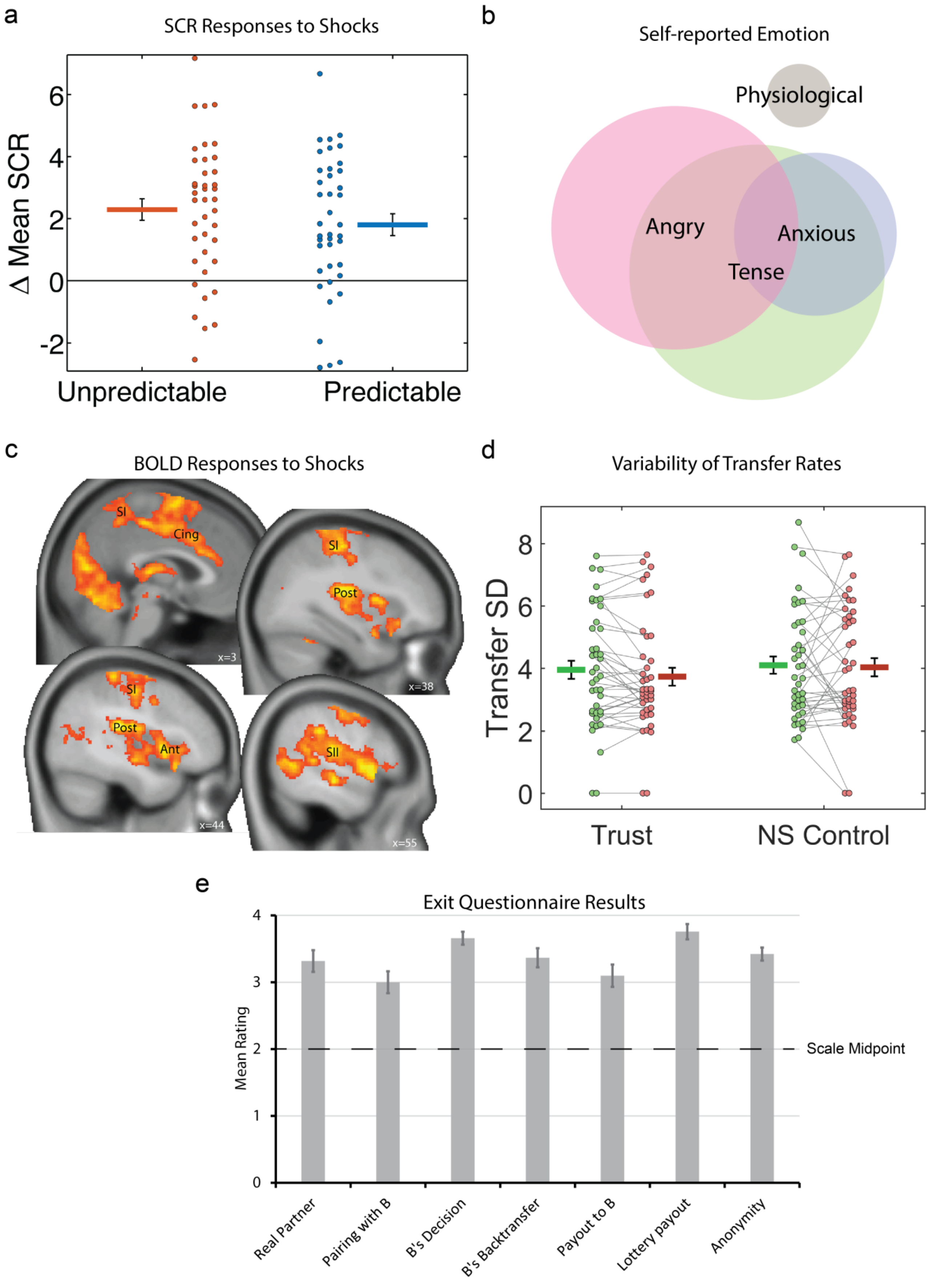

Fig. S1. Manipulation checks. (a) SCR response increases to mildly painful (versus weak) electrical stimulation when the stimulation was unpredictable (orange) and predictable (blue). SCR response to *actually experienced* painful tactile stimulation was significantly increased compared to the experience of a weak tactile stimulation, for both unpredictable [mean difference = 2.29, t(39) = 6.67, $p < 0.0001$] and predictable shock events [mean difference = 1.80, t(39) = 5.15, $p < 0.0001$]. This difference between strong and weak shocks was larger for unpredictable compared to predictable shocks [interaction mean = 0.49, t(39) = 2.16, p = 0.037]. These results, confirming prior research (*62*), show that affective arousal is influenced by complex interactions between cognitive processes such as attention and the predictability of an aversive event.

(b)Self-reported experience of emotion during threat blocks. Venn diagram illustrating proportions and overlap between self-reported emotional reactions to painful tactile stimulation. Note that subjects were able to report multiple emotions in response to the open-ended question about how they felt during threat blocks in an exit questionnaire. 95.12% of subjects indicated that they experienced aversive emotional arousal during pain blocks, such as tenseness and/or anger and/or anxiety. To illustrate the frequency of emotional reactions to painful tactile stimulation, subjects' answers were binned into three emotion categories that best summarize their emotional responses: angry and annoyed responses were grouped into the "Angry" category (25), tense, stressed and surprised responses were grouped into the "Tense" category (27), and scared, nervous, helpless, and sad responses were grouped into the "Anxious" category (11). Most subjects reported angry (61%) and tense (66%) emotions while a minority also felt anxious (27%). A fourth category was termed "physiological" and best characterizes responses that reported physiological reactions related to aversive emotional arousal during painful blocks, such as "sweating" (7%). Of note, due to the open-ended nature of the questions, subjects were able to indicate more than one emotional reaction, as illustrated by the overlap between different sets in the Venn diagram.

(c) Administration of painful compared to just-noticeable tactile stimulation led to increased activity within key regions of the pain matrix, including primary (SI) and secondary (SII) somatosensory cortex, anterior (Ant) and posterior (Post) insula, as well as mid cingulate cortex (Cing). The contrast images reflect activation at the time of strong vs. weak tactile stimulation and are thresholded at $p < 0.05$ FWE-corrected at the cluster level (with a cluster-forming voxel-level threshold of $p<0.001$).

(d) To address whether decisions became less systematic under threat of shock, we investigated choice inconsistencies across the threat and no threat conditions by computing the standard deviation of each participant's transfer amount in the trust and the non-social control game under both threat and no threat. Our logic here is that an increase in the variability of transfer amounts serves as a proxy of erratic decision-making and decision inconsistencies. Therefore, if we find increased variability in the transfer amounts in either the risk or the trust game, this would be reflective of increasingly impulsive decisions, or of not considering all the available choice options (Keinan, 1987). We entered the standard deviation of transfer amounts into a 2-way repeated measures ANOVA with the factors game type and threat and do not find any significant main effects [game type: F(1,40) = 1.736, p = 0.195, $\eta^2 = 0.042$; threat: F(1,40) = 1.333, p = 0.255, $\eta^2 = 0.032$], nor a significant interaction, F(1, 40) = 0.251, p = 0.619, $\eta^2 = 0.006$. These results indicate that choice variability did not differ across threat conditions and game types.

(e) We assessed subjects' believes about interacting with another person and their perception of the social nature of their interactions during the main experiment in an exit questionnaire. The following questions specifically assessed subjects' experience of the social nature of the delayed trust game interactions:

Q1. "In all trust games, I was playing with a real person that was not physically present."
Q2. "In every new trust game I was randomly paired with another participant B."
Q3. "The decisions of participants B were collected in a separate experiment."
Q4. "My earnings in the trust game were based on decisions that other participants made."
Q5. "My decisions in the payout-relevant trial will be transferred to the relevant participant B."

Subjects reported an average (standard error) of 3.29 (0.15) on these five questions, which is significantly larger than the midpoint of the scale (2): Z = 8.499, $p < 0.0001$. This indicates that, throughout the experiment, subjects believed that they were interacting with a real partner who was determining their backtransfer amounts and that our subjects' decisions also determined additional payouts for their interaction partner. For completeness, we also list the additional statements that were included in the exit questionnaire:

Q6. "My payouts in the risk game are based on the result of a computer algorithm."
Q7. "My decisions, as well as these of the other participant, are being treated anonymously."

The average answers and their standard errors are shown in fig. S1e. All answers are significantly greater than the midpoint of the scale demarcated by 2 (all $p < 0.0001$), indicating that subjects believed and trusted our experimental instructions.

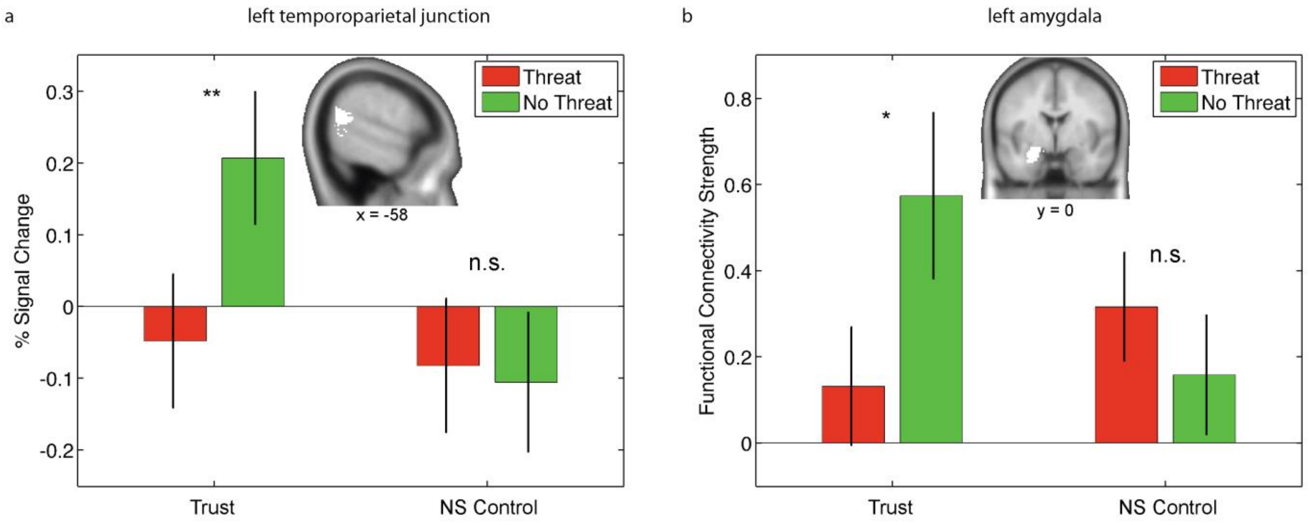

Fig. S2. Additional post hoc inspection of the significant interactions reported in the main paper within all voxels of independent TPJ and amygdala masks. To characterize the interaction patterns reported in the main paper, we extracted subject-specific regression coefficients (beta weights) from all voxels within the entire left TPJ (k = 1031) and amygdala (k = 439) masks and performed post-hoc statistical analyses in independent ROIs. (a) Post-hoc pairwise comparisons in the TPJ ROI revealed significantly greater choice-related activation during no-threat relative to threat when subjects made decisions in the trust task [t(40) = 2.769, p = 0.0085], but not when they made decisions in the control task [t(40) = - 0.167, p = 0.8682]. (b) Post-hoc pairwise comparisons in the amygdala ROI revealed significantly greater choice-related functional connectivity during no-threat relative to threat when subjects made decisions in the trust task [t(40) = 2.097, p = 0.0424], but not when they made decisions in the control task [t(40) = - 0.812, p = 0.4216]. These results confirm that the presence of threat during trust decisions led to a significant suppression of (a) TPJ activity and (b) TPJ-amygdala connectivity relative to the absence of threat.

** p < 0.01; * p < 0.05; n.s.: not significant

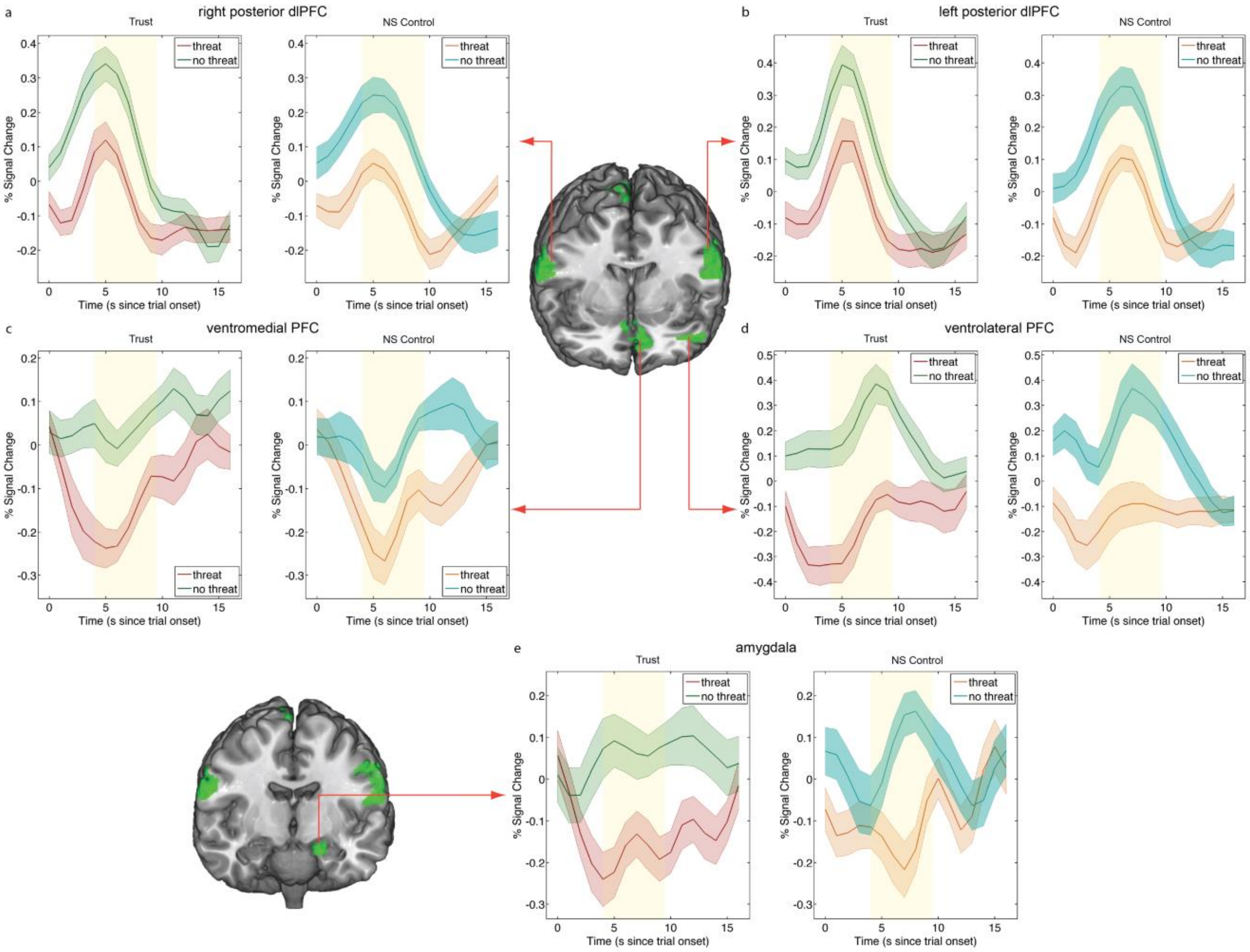

Fig. S3. Main effect of threat 1: Suppression of game type–independent neural correlates of decision-making. We tested the main effect of threat on the neural correlates of decision-making independent of the game type (by including both trust and non-social decisions). This analysis revealed a domain-general network of regions showing significant threat-related reduction (no threat > threat) in choice-related activity during both trust and non-social control trials (see Supplementary Table 7a), consisting of bilateral posterior dlPFC [(a) right: 62, -6, 28, k = 1010 and (b) left: -62, -4, 18, k = 1901], left amygdala [(f) -24, -15, -23, k = 552], posterior paracentral lobule [not shown, 4 -36, 69, k = 887], and a large cluster in ventral anterior prefrontal cortex (k = 4082) that includes vmPFC [(c) -10, 44, -8] and left vlPFC [(d) -48, 41, -8]. Time courses illustrate suppression (a-e) due to threat during trust and non-social control decisions and are shown for both trust (left, threat shown in red, no threat shown in green) and control (right, threat shown in orange, no threat shown in aqua green) decisions in separate graphs. The 5.5-second choice period is displayed in yellow and is shifted 4-seconds to account for the hemodynamic lag.

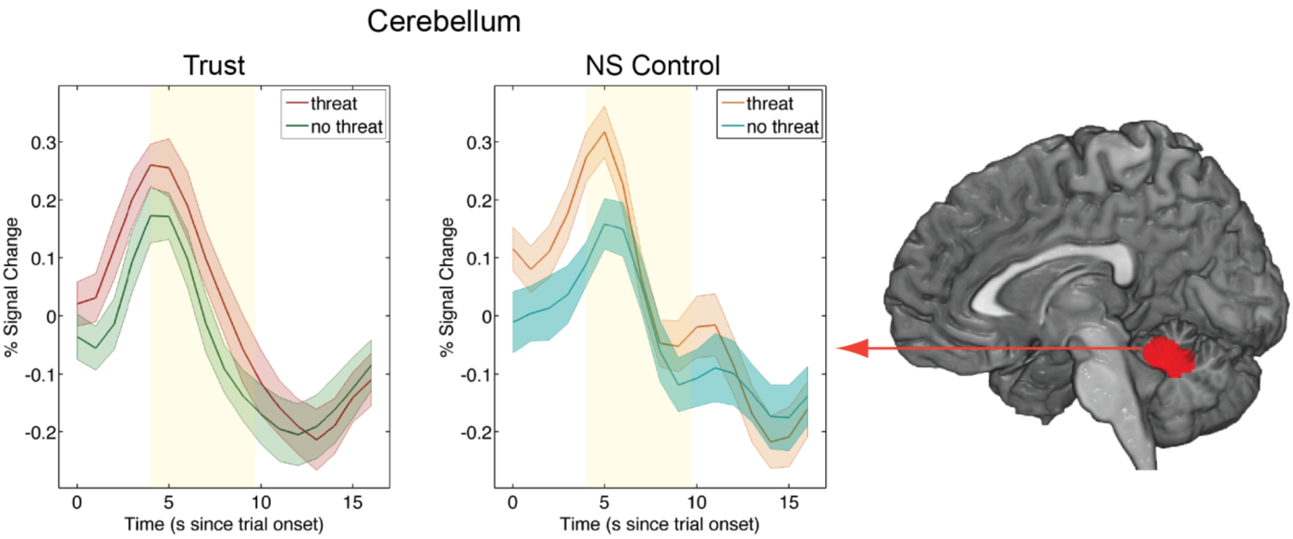

Fig. S4. Main effect of threat 2: Enhancement of game type–independent neural correlates of decision-making. We tested the main effect of threat on the neural correlates of decision-making independent of the game type (including both trust and non-social decisions). In addition to regions showing significant suppression of choice-related activity (fig. S4, Supplementary Table 7a), we also found significant threat-related enhancement of activity (threat > no threat) during the decision phase (see Supplementary Table 7b) in the thalamus [(a) 18, -6, 1, k = 559] and cerebellum [(b) -4, -46, -24, k = 849]. Time courses illustrate enhancements (a-b) due to threat during trust and non-social control decisions and are shown for both trust (left, threat shown in red, no threat shown in green) and NS control (right, threat shown in orange, no threat shown in aqua green) decisions in separate graphs. The 5.5-second choice period is displayed in yellow and is shifted 4-seconds to account for the hemodynamic lag.

3 Supplementary Tables

Table S2. ROI analyses investigating trust-specific neural correlates in regions associated with social cognition and valuation. (see Fig. 2). (a) Regions that are preferentially involved in the trust game relative to the non-social control task when threat is absent. (b) Only the temporoparietal junction (TPJ) shows a significant interaction effect reflective of a change in the neural impact of threat during trust taking relative to the non-social control task. (c) Regions that show significant simple effects of threat during trust decisions.

Structure	L/R	Cluster Size	Peak t	x	y	z	Peak p
(a) Main effect: Trust (No Threat + Threat) > NS Control (No Threat + Threat)							
TPJ	left	9	3.89	-57	-60	27	0.015
(b) Interaction: Trust (No Threat > Threat) – NS Control (No Threat > Threat)							
TPJ	left	34	3.63	-60	-54	19	0.035
Anterior Insula	left	21	4.1	-46	14	-12	0.022
(c) Simple effect: Trust (No Threat > Threat)							
TPJ	left	35	3.97	-58	-55	19	0.012
vmPFC	bil.	227	4.66	9	33	-12	0.002
Ventral striatum	left	19	4.09	-3	3	-6	0.010

Activation clusters survive small volume correction for multiple comparisons based on FWE correction at the peak level for a priori regions of interest using masks created with reverse inference maps from relevant search terms on the neurosynth.org database. The left TPJ masks contained 1031 voxels; the dorsomedial and ventromedial PFC masks contained 3175 and 1327 voxels, respectively, the left ventral striatum mask contained 3239 voxels, and the anterior insula mask contained 13844 voxels.

Table S3. ROI analyses investigating TPJ-amygdala connectivity patterns. (see Fig. 3c-d)**.** (a) TPJ-amygdala connectivity during decision-making in the trust relative to the control task *when threat is absent* and (b-c) the differential impact of threat on TPJ-amygdala connectivity in the trust task relative to the control task.

Structure	**L/R**	**Cluster Size**	**Peak t**	**x**	**y**	**z**	**Peak p**
(a) Main effect: No Threat (Trust + NS Control) > Threat (Trust + NS Control)							
amygdala	left	14	3.78	-28	-6	-14	0.008
(b) Interaction: Trust (No Threat > Threat) – NS Control (No Threat > Threat)							
amygdala	left	12	3.68	-26	0	-23	0.012
(c) Simple effect: Trust (No Threat > Threat)							
amygdala	left	23	4.03	-28	-6	-14	0.004
amygdala	left	3	3.39	-26	0	-23	0.027
(d) Simple effect: Threat (NS Control > Trust)							
amygdala	left	14	3.24	-26	0	-23	0.035

Activation clusters survive small volume correction for multiple comparisons based on FWE correction at the peak level for a priori regions of interest using anatomical masks created with reverse inference maps from relevant search terms on the neurosynth.org database. The left amygdala mask contained 439 voxels.

Table S4. ROI analyses investigating brain-behavior relationships. ROI analyses investigating how the correlation between TPJ functional connectivity with its target regions and mean transfer rate is influenced by (a) game type in the absence of threat; (b) threat in the trust but not the NS control game; and (c) threat during trust decisions.

Structure	L/R	Cluster Size	Peak t	x	y	z	Peak p
(a) Main effect: Trust (No Threat + Threat) > NS Control (No Threat + Threat)							
STS (TPJ)	right	30	3.75	64	-43	4	0.034
STS (TPJ)	left	10	4.59	-62	-52	-5	0.001
DMPFC	bil.	30	4.04	-12	54	40	0.021
DMPFC	bil	9	3.91	-3	38	48	0.037
Amygdala	right	16	3.49	28	2	-20	0.024
Anterior Insula	left	245	5.16	-51	21	-6	<0.001
Anterior Insula	right	525	4.5	56	18	1	0.006
(b) Interaction: Trust (No Threat > Threat) – NS Control (No Threat > Threat)							
STS (TPJ)	right	7	4.73	64	-43	6	0.001
(c) Simple effect: Trust (No Threat > Threat)							
STS (TPJ)	right	31	5.63	64	-43	6	<0.001
Anterior Insula	Left	86	3.95	-46	23	-8	0.035

Activation clusters survive small volume correction for multiple comparisons based on FWE correction at the peak level for a priori regions of interest using masks created with reverse inference maps from relevant search terms on the neurosynth.org database. The right amygdala mask contained 492 voxels, the right TPJ mask contained 1416 voxels, the dmPFC mask contained 3175 voxels, the left anterior insula mask contained 13844 voxels, the right anterior insula mask contained 13871 voxels.

Table S5. Whole-brain analyses investigating brain-behavior relationships. (see Fig. 3). Results from whole-brain analysis investigating how the correlation between TPJ functional connectivity with its target regions and mean transfer rate is influenced by (a) game type in the absence of threat; (b) threat in the trust but not the NS control game; and (c) threat during trust decisions ($p < 0.001$, cluster size > 226, FWE corrected at cluster-level).

Structure	L/R	Cluster Size	Peak t	x	y	z
(a) Main effect: Trust (No Threat + Threat) > NS Control (No Threat + Threat)						
DMPFC / SMA	bil.	1721	5.82	-2	17	54
Inferior frontal gyrus (IFG)	left	343	5.16	-51	21	-6
posterior insula	right	310	4.79	38	-13	12
dmPFC / superior frontal gyrus (SFG)	left	416	4.78	-16	53	40
IFG / anterior insula	right	1422	4.62	42	27	-11
DLPFC / MFG	right	601	4.56	44	8	28
DLPFC / MFG	left	415	4.45	-52	6	18
Superior frontal gyrus (SFG)	right	253	4.17	26	0	61
(b) Interaction: Trust (No Threat > Threat) – NS Control (No Threat > Threat)						
*Superior Temporal Sulcus**	*right*	*54*	*4.73*	*64*	*-43*	*6*
(c) Simple effect: Trust (No Threat > Threat)						
Superior Temporal Sulcus	right	280	5.63	64	-43	6
*Cuneus**	*right*	*204*	*5.51*	*18*	*-82*	*16*
(d) Simple effect: Trust (No Threat)						
Superior Temporal Sulcus	right	335	5.95	64	-43	4
Intraparietal Sulcus (IPS)	right	691	5.17	44	-48	40
Superior frontal gyrus (SFG)	right	475	4.89	20	50	39
Inferior frontal gyrus (IFG)	left	429	4.87	-51	27	-3
Posterior insula	left	578	4.59	-40	3	7
DLPFC, MFG	right	1096	4.43	38	27	40
DMPFC, SMA	bilateral	667	4.42	-2	26	40
Precuneus	right	216	3.9	28	-63	56

*survives whole-brain FWE-correction at initial cluster-forming threshold of $p < 0.005$

Table S6. Whole-brain analyses investigating the main effect of threat. (see fig. S4). Results from whole-brain analysis investigating the main effect of threat on neural activity during decision-making independent of game type ($p < 0.001$, cluster size > 226, FWE corrected at cluster-level).

Structure	**L/R**	**Cluster Size**	**Peak t**	**x**	**y**	**z**
(a) Main Effect: no threat > threat						
Frontopolar cortex	bil.	646	5.91	-14	68	3
Inferior frontal gyrus / ventrolateral PFC	Left	281	5.59	-48	41	-8
Ventromedial PFC	Bil.	464	4.77	-10	44	-8
Ventromedial PFC	Bil.	316	4.74	6	21	-14
posterior dorsolateral PFC	right	515	4.67	62	-6	28
posterior dorsolateral PFC	left	686	4.49	-62	-4	18
posterior paracentral lobule	bil.	309	4.36	4	-36	69
Amygdala	Left	226	4.07	-24	-15	-23
(b) Main Effect: threat > no threat						
cerebellum	bil.	549	-2.61	-4	-46	-24
*Thalamus**	*right*	*133*	*-2.61*	*6*	*-7*	*4*