nature neuroscience

Corresponding Author: Tor Wager

Manuscript Number: NN-A46884A

Manuscript Type: Article

Main Figures: 4

Supplementary Figures: 10

Supplementary Tables: 3

Supplementary Videos: 0

Reporting Checklist for Nature Neuroscience

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. For more information, please read Reporting Life Sciences Research.

Please note that in the event of publication, it is mandatory that authors include all relevant methodological and statistical information in the manuscript.

Statistics reporting, by figure

- Please specify the following information for each panel reporting quantitative data, and where each item is reported (section, e.g. Results, & paragraph number).
- Each figure legend should ideally contain an exact sample size (n) for each experimental group/condition, where n is an exact number and not a range, a clear definition of how n is defined (for example x cells from x slices from x animals from x litters, collected over x days), a description of the statistical test used, the results of the tests, any descriptive statistics and clearly defined error bars if applicable.
- For any experiments using custom statistics, please indicate the test used and stats obtained for each experiment.
- Each figure legend should include a statement of how many times the experiment shown was replicated in the lab; the details of sample collection should be sufficiently clear so that the replicability of the experiment is obvious to the reader.
- For experiments reported in the text but not in the figures, please use the paragraph number instead of the figure number.

Note: Mean and standard deviation are not appropriate on small samples, and plotting independent data points is usually more informative. When technical replicates are reported, error and significance measures reflect the experimental variability and not the variability of the biological process; it is misleading not to state this clearly.

		TEST USED		n			DESCRIPTIVE STATS (AVERAGE, VARIANCE)		P VALUE		DEGREES OF FREEDOM & F/t/z/R/ETC VALUE	
	FIGURE NUMBER	WHICH TEST?	SECTION & PARAGRAPH #	EXACT VALUE	DEFINED?	SECTION & PARAGRAPH #	REPORTED?	SECTION & PARAGRAPH #	EXACT VALUE	SECTION & PARAGRAPH #	VALUE	SECTION & PARAGRAPH #
example	1a	one-way ANOVA	Fig. legend	9, 9, 10, 15	mice from at least 3 litters/group	Methods para 8	error bars are mean +/- SEM	Fig. legend	p = 0.044	Fig. legend	F(3, 36) = 2.97	Fig. legend
example	results, para 6	unpaired t-test	Results para 6	15	slices from 10 mice	Results para 6	error bars are mean +/- SEM	Results para 6	p = 0.0006	Results para 6	t(28) = 2.808	Results para 6

		TEST USED		n			DESCRIPTIVE STATS (AVERAGE, VARIANCE)		P VALUE		DEGREES OF FREEDOM & F/t/z/R/ETC VALUE	
	FIGURE NUMBER	WHICH TEST?	SECTION & PARAGRAPH #	EXACT VALUE	DEFINED?	SECTION & PARAGRAPH #	REPORTED?	SECTION & PARAGRAPH #	EXACT VALUE	SECTION & PARAGRAPH #	VALUE	SECTION & PARAGRAPH #
+ -	1c	multi-level logistic regressions	results -> behavioral results, para 1; Online Methods -> Logistic regression model; Fig. legend	23	yes	Fig. legend, Online Methods -> participants, para 1	in figure, not in legend	n/a	p < 0.05 , see figure legend	results-> behavioral results, para 1;	one trial back (t(22) = 9.20, p<0.001) and two trials back (t(22) = 4.36, p<0.001)	Fig. legend
+ -	1d	1st level: parametric modulation 2nd level: robust group-level t-tests	results -> Correlations with model-based aversive prediction errors, para 1, Online methods -> fMRI data analyses, para 1	23	yes	Fig. legend, Online Methods -> participants, para 1	no, but does not really apply to imaging data	n/a	P < .05, FWER corrected based on cluster extent (cluster-defining thresholds of 0.001, 0.01 and 0.05	Online methods -> fMRI data analyses -> model---based PE analysis; Figure and Figure legend	multiple tests - see supplementary table 1	multiple tests - see supplementary table 1

+ -	2b	slope sign permutation test on data extracted per quartile of expected value in NAcc and PAG ROIs	Results -> Axiomatic tests of aversive prediction errors, para 4; Online methods -> fMRI data analyses -> ROI axiomatic response profile analysis, para 3	23	yes	Fig. legend, Online Methods -> participants, para 1	in figure, not in legend	n/a	p < 0.05	Results -> Axiomatic tests of aversive prediction errors, para 4;	Axiom #1; t(22) = 3.67, p <0.05). Axiom #2; Pain trials: t(22)=-2.05, p < 0.05; No-Stimulus trials: t(22)=-1.98, p < 0.05).	Results -> Axiomatic tests of aversive prediction errors, para 4;
+ -	2b	t-tests in NAcc and PAG ROIs, completed by Bayesian analyses testing the null	Results -> Axiomatic tests of aversive prediction errors, para 4; Online methods -> fMRI data analyses -> ROI axiomatic response profile analysis, para 3	23	yes	Fig. legend, Online Methods -> participants, para 1	in figure, not in legend	n/a	p < 0.05, log odds in favour of the null reported in text	Results -> Axiomatic tests of aversive prediction errors, para 4;	Axiom #3; t(22) = 0.13, p = n.s.	Results -> Axiomatic tests of aversive prediction errors, para 4;

+ -	3a	1st level: parametric modulation 2nd level: robust group-level t-tests and conjunction	Results -> Axiomatic tests of aversive prediction errors, para 5; Online methods -> fMRI data analyses -> conjunction analysis	23	yes	Fig. legend, Online Methods -> participants, para 1	no, but does not really apply to imaging data	n/a	P < .05 (one-tailed), FWER corrected based on cluster extent (cluster-defining threshold of 0.05)	Online methods -> fMRI data analyses -> conjunction analysis; Figure and Figure legend	multiple tests - see supplementary table 2	supplementray table 2
+ -	4ab	retained DCM model	Results -> Network dynamics underlying the generation of aversive PE signals Online Methods -> fMRI data analyses -> Dynamic causal models.	23	yes	Fig. legend, Online Methods -> participants, para 1	n/a	n/a	n/a	n/a	n/a	n/a
+ -	S1	descriptive	results -> Correlations with model-based aversive prediction errors	23	yes	Fig. legend, Online Methods -> participants, para 1	in figure, not in legend	n/a	n/a	n/a	n/a	n/a

+ -	S2c	t-test against the null hypothesis that ROI activity is equal to 0	Results -> Comparing aversive and reward PEs in the PAG vs. VS	pain = 23, money = 21	yes	Fig. legend, Online Methods -> Study 2 -> participants, para 1	in figure, not in legend	n/a	VS. p < 0.001 PAG, p = 0.14	Results -> Comparing aversive and reward PEs in the PAG vs. VS	dl_pain = 22 dl_money = 20 PAG-pain; t = 3.07 PAG-money; t = 1.54; VS-pain; t = 0.79; VS-money = t = 5.7	Results -> Comparing aversive and reward PEs in the PAG vs. VS
+ -	S3e	ANOVA	Results -> Replication of aversive PEs and extension to varying levels of pain	50	yes	Fig. legend, Online Methods -> Study 3 -> participants, para 1	in figure, not in legend	n/a	cue effect: p < 0.05 instruction effects: p < 0.001	Results -> Replication of aversive PEs and extension to varying levels of pain	cue effect: F = 4.39 instruction effects: F = 16.03	Results -> Replication of aversive PEs and extension to varying levels of pain
+ -	S4	1st level: parametric modulation 2nd level: robust group-level t-tests;	Online methods -> fMRI data analyses -> model-based PE analysis. Results -> Expectancy effects at the time of decision Results -> expected probability of avoidance	23	yes	Fig. legend, Online Methods -> Study 3 -> participants, para 1	no, but does not really apply to imaging data	n/a	P < .05, FWER corrected based on cluster extent (cluster-defining thresholds of 0.001, 0.01 and 0.05)	Online methods -> fMRI data analyses -> model-based PE analysis. Figure and figure legend	dl:22 multiple tests - see supplementary table 3	supplementary table 3

+ -	S5	1st level: parametric modulation 2nd level: robust group-level t-tests; conjunction	results -> Correlations with model-based aversive prediction errors, para 1, Online methods -> fMRI data analyses, para 1	23	yes	Fig. legend, Online Methods -> Study 3 -> participants, para 1	no, but does not really apply to imaging data	n/a	P < .05, FWER corrected based on cluster extent (cluster-defining thresholds of 0.001, 0.01 and 0.05)	results -> Correlations with model-based aversive prediction errors, para 1, Online methods -> fMRI data analyses, para 1	dl:22 multiple tests - see supplementary table 1	supplementary table 1
+ -	S6	no test/task design	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
+ -	S7b-10b	Bayesian model selection process: expected and exceedence probability	Results -> Network dynamics underlying the generation of aversive PE signals Online Methods -> fMRI data analyses -> Dynamic causal models	23	yes	Fig. legend, Online Methods -> Study 3 -> participants, para 1	n/a	n/a	n/a	n/a	n/a	n/a

Representative figures

1. Are any representative images shown (including Western blots and immunohistochemistry/staining) in the paper?

   If so, what figure(s)?

   Figure 1b is example data from one representative subject.

2. For each representative image, is there a clear statement of how many times this experiment was successfully repeated and a discussion of any limitations in repeatability?

   If so, where is this reported (section, paragraph #)?

   N/A; this figure is just to illustrate task design.

Statistics and general methods

1. Is there a justification of the sample size?

   If so, how was it justified?

   Where (section, paragraph #)?

   Even if no sample size calculation was performed, authors should report why the sample size is adequate to measure their effect size.

   We used standard sample sizes (Study 1, N=23; Study 2, N=21, Study 3, N=50) for fMRI experiments based on power calculations to detect moderately large effects. We did not specifically justify this sample size in the manuscript.

2. Are statistical tests justified as appropriate for every figure?

   Where (section, paragraph #)?

   All the statistical tests are explained in the results section , online methods, figure captions for study #1. For whole-brain analyses, we use mainly standard robust group-level t-tests across general linear model parameter estimates, treating subject as a random effect. For ROI axiomatic tests, we use a combination of t-tests (axiom #1), slope sign permutation tests (axiom #2), and bayesian analyses of odds in favour/against the null (axiom #3). For other ROI analyses, we use ANOVAs (Study #2 and #3), and t-tests (study #2).

   a. If there is a section summarizing the statistical methods in the methods, is the statistical test for each experiment clearly defined?

      Yes, all the statistical tests are explained in the results sections, online methods, figure captions, and supplementary materials, as described above.

   b. Do the data meet the assumptions of the specific statistical test you chose (e.g. normality for a parametric test)?

      Where is this described (section, paragraph #)?

      Multi-level logistic regression results for behavior data and significant brain findings were checked for outliers, normality, and homoskedasticity. We do not specifically report this assessment in the manuscript, but could include it if deemed important. Robust regression results are designed to be resistant to outliers and violations of normality. Sign permutation tests employ a well-established bootstrapping procedure that does not assume normality.

   c. Is there any estimate of variance within each group of data?

      Is the variance similar between groups that are being statistically compared?

      Where is this described (section, paragraph #)?

      We have only one group for each of our studies. Therefore all of our manipulations were within-subject.

   d. Are tests specified as one- or two-sided?

      All the P-values reported in this paper are two-tailed (two-sided test), including brain results. The only exception is the whole-brain conjunction search presented in Figure 3a. The explicit directionality of the axioms warranted the use of one-tailed tests there.

   e. Are there adjustments for multiple comparisons?

      All fMRI results were corrected for multiple comparisons at familywise error rate P < .05 corrected, based on cluster extent (thresholds reported in the paper).

3. Are criteria for excluding data points reported?

   Was this criterion established prior to data collection?

   Where is this described (section, paragraph #)?

   Data from 3 participants were excluded based on their bad performance on the task. This is a fairly common procedure for this type of task (see Online methods -> Participants, paragraph #1 and Online methods -> reinforcement model-based analysis, paragraph #2)

4. Define the method of randomization used to assign subjects (or samples) to the experimental groups and to collect and process data. If no randomization was used, state so. Where does this appear (section, paragraph #)?	Because we have only one group per study, there is no randomization of group assignment. However, in study #1, each participant was randomly assigned one of four sets of random walks governing reinforcement probabilities (Online methods -> experimental task).
5. Is a statement of the extent to which investigator knew the group allocation during the experiment and in assessing outcome included? If no blinding was done, state so. Where (section, paragraph #)?	We have only one group per study, so this does not apply.
6. For experiments in live vertebrates, is a statement of compliance with ethical guidelines/regulations included? Where (section, paragraph #)?	n/a
7. Is the species of the animals used reported? Where (section, paragraph #)?	n/a (humans)
8. Is the strain of the animals (including background strains of KO/ transgenic animals used) reported? Where (section, paragraph #)?	n/a
9. Is the sex of the animals/subjects used reported? Where (section, paragraph #)?	n/a
10. Is the age of the animals/subjects reported? Where (section, paragraph #)?	n/a
11. For animals housed in a vivarium, is the light/dark cycle reported? Where (section, paragraph #)?	n/a
12. For animals housed in a vivarium, is the housing group (i.e. number of animals per cage) reported? Where (section, paragraph #)?	n/a
13. For behavioral experiments, is the time of day reported (e.g. light or dark cycle)? Where (section, paragraph #)?	n/a
14. Is the previous history of the animals/subjects (e.g. prior drug administration, surgery, behavioral testing) reported? Where (section, paragraph #)?	n/a

a. If multiple behavioral tests were conducted in the same group of animals, is this reported?

Where (section, paragraph #)?

n/a

15. If any animals/subjects were excluded from analysis, is this reported?

Where (section, paragraph #)?

n/a

a. How were the criteria for exclusion defined?

Where is this described (section, paragraph #)?

n/a

b. Specify reasons for any discrepancy between the number of animals at the beginning and end of the study.

Where is this described (section, paragraph #)?

n/a

Reagents

1. Have antibodies been validated for use in the system under study (assay and species)?

n/a

a. Is antibody catalog number given?

Where does this appear (section, paragraph #)?

n/a

b. Where were the validation data reported (citation, supplementary information, Antibodypedia)?

Where does this appear (section, paragraph #)?

n/a

2. If cell lines were used to reflect the properties of a particular tissue or disease state, is their source identified?

Where (section, paragraph #)?

n/a

a. Were they recently authenticated?

Where is this information reported (section, paragraph #)?

n/a

Data deposition

Data deposition in a public repository is mandatory for:

a. Protein, DNA and RNA sequences
b. Macromolecular structures
c. Crystallographic data for small molecules
d. Microarray data

Deposition is strongly recommended for many other datasets for which structured public repositories exist; more details on our data policy are available here. We encourage the provision of other source data in supplementary information or in unstructured repositories such as Figshare and Dryad.

1. Are accession codes for deposit dates provided? Where (section, paragraph #)?	n/a

Computer code/software

Any custom algorithm/software that is central to the methods must be supplied by the authors in a usable and readable form for readers at the time of publication. However, referees may ask for this information at any time during the review process.

1. Identify all custom software or scripts that were required to conduct the study and where in the procedures each was used.	We used custom scripts for slope sign permutation tests (figure 2 and 3). All custom scripts are freely available (http://wagerlab.colorado.edu/tools)
2. Is computer source code/software provided with the paper or deposited in a public repository? Indicate in what form this is provided or how it can be obtained.	All custom scripts are freely available (http://wagerlab.colorado.edu/tools)

Human subjects

1. Which IRB approved the protocol? Where is this stated (section, paragraph #)?	The Columbia University IRB approved study 1 protocol (p.20). The New York University IRB approved study 2 protocol (supplementary materials, p. 2). The University of Colorado Boulder IRB approved study 3 protocol (supplementary materials, p. 4)
2. Is demographic information on all subjects provided? Where (section, paragraph #)?	yes(online methods -> participants, paragraph #1)
3. Is the number of human subjects, their age and sex clearly defined? Where (section, paragraph #)?	yes(online methods -> participants, paragraph #1)
4. Are the inclusion and exclusion criteria (if any) clearly specified? Where (section, paragraph #)?	yes(online methods -> participants, paragraph #1)
5. How well were the groups matched? Where is this information described (section, paragraph #)?	n/a

6. Is a statement included confirming that informed consent was obtained from all subjects? Where (section, paragraph #)?	yes(online methods -> participants, paragraph #1)
7. For publication of patient photos, is a statement included confirming that consent to publish was obtained? Where (section, paragraph #)?	n/a

fMRI studies

For papers reporting functional imaging (fMRI) results please ensure that these minimal reporting guidelines are met and that all this information is clearly provided in the methods:

1. Were any subjects scanned but then rejected for the analysis after the data was collected?	yes
a. If yes, is the number rejected and reasons for rejection described? Where (section, paragraph #)?	Data from 3 participants were excluded based on their bad performance on the task. This is a fairly common procedure for this type of task (see Online methods -> Participants, paragraph #1 and Online methods -> reinforcement model-based analysis, paragraph #2)
2. Is the number of blocks, trials or experimental units per session and/ or subjects specified? Where (section, paragraph #)?	Yes. Study 1: Online methods ->Experimental task(paragraph 1). Study 2: Online methods -> Study -> Monetary reward task (paragraph 1). Study 3: Online methods -> Study -> Experimental task (paragraph 1).
3. Is the length of each trial and interval between trials specified?	Yes
4. Is a blocked, event-related, or mixed design being used? If applicable, please specify the block length or how the event-related or mixed design was optimized.	
5. Is the task design clearly described? Where (section, paragraph #)?	A slow event-related design was used for study #1, with trials separated by approx. 1 min, to avoid physiological sensitization and habituation to painful events. Therefore, optimization algorithms for rapid event-related designs (e.g., Wager and Nichols, 2003) are not appropriate in this context.
6. How was behavioral performance measured?	For study #1, participants choices were analyzed as a function of reinforcement history with logistic regressions or computational learning models.
7. Is an ANOVA or factorial design being used?	Yes, an ANOVA is used on data extracted from ROIs in study #2 and study #3.

8. For data acquisition, is a whole brain scan used? If not, state area of acquisition.	For study #1, the acquisition covers most of the brain with the exception of dorso-posterior parietal areas.
a. How was this region determined?	We had a priori hypotheses about the ventromedial prefrontal cortex and chose to optimize signal there by thinning slices, resulting in a loss of coverage in parietal cortex.
9. Is the field strength (in Tesla) of the MRI system stated?	yes
a. Is the pulse sequence type (gradient/spin echo, EPI/spiral) stated?	yes
b. Are the field-of-view, matrix size, slice thickness, and TE/TR/ flip angle clearly stated?	yes
10. Are the software and specific parameters (model/functions, smoothing kernel size if applicable, etc.) used for data processing and pre-processing clearly stated?	yes
11. Is the coordinate space for the anatomical/functional imaging data clearly defined as subject/native space or standardized stereotaxic space, e.g., original Talairach, MNI305, ICBM152, etc? Where (section, paragraph #)?	yes (MNI), in Online Methods ->fMRI data acquisition and preprocessing -> preprocessing, paragraph #2.
12. If there was data normalization/standardization to a specific space template, are the type of transformation (linear vs. nonlinear) used and image types being transformed clearly described? Where (section, paragraph #)?	yes, in Online Methods ->fMRI data acquisition and preprocessing -> preprocessing, paragraph #2.
13. How were anatomical locations determined, e.g., via an automated labeling algorithm (AAL), standardized coordinate database (Talairach daemon), probabilistic atlases, etc.?	We visualized significant results on the group-average T1 (anatomical) image and compared results with atlases (e.g., Mai et al., 2004). In addition, we corroborated names of regions with multiple electronic atlases including the Havard-oxford atlas, Automated Anatomical Labeling (AAL) atlas, and Neurosynth.
14. Were any additional regressors (behavioral covariates, motion etc) used?	yes, this is described in Online Methods ->fMRI data acquisition and preprocessing -> preprocessing, paragraph #1.
15. Is the contrast construction clearly defined?	yes, in Online Methods ->fMRI data analyses
16. Is a mixed/random effects or fixed inference used?	We used mixed effects models for all analyses, treating subject as a random effect.
a. If fixed effects inference used, is this justified?	
17. Were repeated measures used (multiple measurements per subject)?	yes
a. If so, are the method to account for within subject correlation and the assumptions made about variance clearly stated?	Yes. We used a standard two-level analyses framework. We report t-tests on planned within-subjects contrasts, and thus do not use repeated-measures methods that require estimation of the inter-subject covariance matrix and whitening (e.g., ANOVA using multiple contrast images at the 2nd level).

18. If the threshold used for inference and visualization in figures varies, is this clearly stated?	Yes. All the threshold methods and levels are stated in each figure legend and in the methods section.
19. Are statistical inferences corrected for multiple comparisons?	yes
a. If not, is this labeled as uncorrected?	
20. Are the results based on an ROI (region of interest) analysis?	yes
a. If so, is the rationale clearly described?	yes
b. How were the ROI's defined (functional vs anatomical localization)?	NAcc ROI was based on a previous paper (Rutledge et al., 2010). PAG ROI was created by positioning 3 6-mm spheres along the aqueduct.
21. Is there correction for multiple comparisons within each voxel?	We used cluster-extent based thresholding to correct for multiple comparisons ($P < .05$ FWER corrected based on spatial extent). This is currently the most widely used multiple comparisons correction procedure in the field.
22. For cluster-wise significance, is the cluster-defining threshold and the corrected significance level defined?	Yes. In addition, details can be found in supplementary tables

Additional comments

Additional Comments	