Title: A novel analytical decoder of BOLD signals for dissociating latent neurobehavioral processes

Authors: Shitong Xiang[1,2†], Tianye Jia[1,2,3,4†*], Chao Xie[1,2†], Zhichao Zhu[1,2], Jujiao Kang[1,2,7], Wei Cheng[1,2], Gunter Schumann[4,5], Trevor W. Robbins[1,2,6*], Jianfeng Feng[1,2,7,8,9*]

Affiliations:

[1] Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China.

[2] Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, China;

[3] Centre for Population Neuroscience and Precision Medicine (PONS), Institute of Psychiatry, Psychology & Neuroscience, SGDP Centre, King's College London, United Kingdom, SE5 8AF;

[4] Centre for Population Neuroscience and Precision Medicine (PONS), Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China.

[5] Centre for Population Neuroscience and Precision Medicine (PONS), Dept. of Psychiatry and Psychotherapy, CCM, Charite Universitaetsmedizin, Berlin, Germany.

[6] Department of Psychology and Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, United Kingdom;

[7] Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China;

[8] Department of Computer Science, University of Warwick, Coventry, United Kingdom;

[9] School of Mathematical Sciences and Centre for Computational Systems Biology, Fudan University, Shanghai, China.

† These authors contributed equally to this work.

* Correspondence authors.

Jianfeng Feng (Address: Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, Shanghai, 200433, China. Email: jianfeng64@gmail.com)

or

Tianye Jia (Address: Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, Shanghai, 200433, China. Email: tianyejia@fudan.edu.cn)

or

Trevor W. Robbins (Address: Department of Psychology and Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, United Kingdom. Email: twr2@cam.ac.uk)

Abstract:

Brain, as a complex cognitive system, often processes multiple dimension information synchronously and integrate them to adapt dynamic environments and make effective decisions.[1-3] How to retrieve latent neurobehavioral processes from complex human neurobiological signals is an important yet previously unresolved challenge.[4,5] For instance, the previous literature has proposed two fundamental yet mutually confounded processes during the decision making and affective processing, i.e. valance and arousal.[6,7] Here, we develop a novel analytical approach, orthogonal-Decoding multi-Cognitive Processes (DeCoP), with which we dissociate neural responses in processing valence and arousal information during tests of motivational and emotional function. During reward/punishment anticipation, we decode brain-wide responses into spatially overlapping, yet functionally independent, evaluation and readiness networks, i.e., motivational valence and arousal processing, which are modulated differentially by the meso-limbic vs nigro-striatal dopamine systems. Similarly, during emotional reactivity, we decompose amygdala response into independent emotional valence and facial arousal processing features. We demonstrate that DeCoP can resolve paradoxically unexpected brain 'inactivation', and be applied more generally to decode multiple latent neurobehavioral processes. Furthermore, we anticipate our approach to advance both the design and hypothesis testing of cognitive experimental task paradigms.

Main

The brain frequently engages parallel processing involving different latent behavioural processes mediated by functionally distinct, though spatially over-lapping, neural networks. Previously, human functional neuroimaging studies have been unsuccessful in unravelling these processes from basal compound physiological signals. It is therefore challenging to build process-specific and mechanistic models of the brain or develop robust biomarkers for dysfunctional processes in psychiatric disorders. Recent overarching frameworks propose two different cognitive processes engaged in parallel during rewarded or punished behaviour, namely evaluation (i.e. valence processing, scaling signal values from reward to punishment) and response readiness (subsuming arousal and attentional salience, contributing to response preparatory processes) [3,6]. The evaluation process is essential for guiding upcoming action selections [8], for which the brain has evolved dedicated regions/circuits for evaluating the value of actions [9-12]. Complementary to evaluation, both reward and punishment, as highly salient events, attract higher attention than neutral stimuli, also engaging greater levels of motor preparation and emotional arousal [13-15], hence contributing to response (including motor) readiness. Therefore, evaluation and readiness signals are inevitably confounded with each other during reward/punishment processing. Unfortunately, decomposing this compound signal, for example in human fMRI studies, has proven challenging because independent components cannot be identified in many experimental test paradigms, although previous attempts have made to overcome this problem [4,5,16]. However, existing approaches have failed to disentangle signals in brain regions known to encode both evaluation and readiness signals, for example, in the striatum and the ventromedial prefrontal cortex (vmPFC) [15], and have not provided convincing evidence to either prove or refute the assumption of independence of the signals. To resolve this complex theoretical issue, we developed a novel approach, orthogonal-Decoding multi-Cognitive Processes (DeCoP), which, for the first time, allows for a brain-wide decomposition of process-specific neurobiological representations of complex neurobehavioral processes. This technique has wide

application for decomposing compound neuroimaging signals, such as the BOLD response.

Experiment and model designs of DeCoP

A monetary incentive delay (MID) task (Fig. 1a), one of the classical and widely used fMRI paradigms for reward processing, was conducted in 1939 children aged 9-10 from the ABCD study (Extended Data Table. 1)[17] to assess reward/punishment processing with gain/loss conditions for small or large amounts of money and a neutral condition (i.e., -5.0 $, -0.2 $, 0, 0.2 $ or 5.0 $). Specifically, in the second-level analyses of the BOLD signal, based on these amounts, we defined plausible orthogonal contrasts [-2, -1, 0, 1, 2] and [2, 1, 0, 1, 2] that reflected putative independent hypothetical processes of evaluation and readiness, respectively (i.e. a large reward and a large punishment are assumed to be equally salient). Their complementary orthogonal contrasts (i.e., [-1, 2, 0, -2, 1] and [1, -2, 2, -2, 1]) could explain information not accounted for by the above hypothetical models (Fig. 1b). We were thus able to retrieve the decomposed signal components of those confounded processes (Extended Data Fig. 1a, see Methods for more details).

Decompose neural signals with DeCoP

The vmPFC (Brodmann area [BA] 10-11; Peak MNI: [-9, 49, -9], t_{1938} = 15.69, Cohen's D = 0.36, $p_{\text{uncorrected}}$ = 2.55E-52) and ventral striatum (VS, Peak MNI: [7, 25, -3], t_{1938} = 14.21, Cohen's D = 0.32, $p_{\text{uncorrected}}$ = 1.14E-43) were the most prominent regions identified in the evaluation model (Fig. 1c upper left&1d), thus being highly sensitive for tracking the entire dimension from punishment to reward. These areas coincide with the terminal regions of the dopamine neuron projections from the ventral tegmental area (VTA), i.e. the meso-corticolimbic dopamine system [8,18,19]. For the readiness model, however, the signals were more widely dispersed across cortical and subcortical areas, including motor-somatosensory, salience and attention networks, and regions such as the dorsal striatum (DS, Peak MNI: [0, 10, 4], t_{1938} = 32.80, Cohen's D = 0.75, $p_{\text{uncorrected}}$ = 4.56E-188) and thalamus (THA, Peak MNI: [13, -6, 16], t_{1938} = 24.66, Cohen's D = 0.56, $p_{\text{uncorrected}}$ = 5.23E-117) (Fig. 1c upper right &1d), consistent

with their engagement in processing both reward and punishment [20].

Neural circuits for decomposed signals

We then investigated whether the neural representations of evaluation and readiness signals were underpinned by different neural circuits, in particular those modulated putatively by the midbrain dopaminergic projections originating from either the substantia nigra pars compacta (SNc) or the VTA, which plays a central role in reward prediction and approach [18,19]. We found regions of the evaluation model with higher functional connectivity (FC) to VTA than to SNc (*paired t*-test: t_{183} = 14.84, Cohen's D = 1.10, p < 10E-32, Fig. 2a), and regions of the readiness model with higher FC to the SNc than to the VTA (*paired t*-test: t_{183} = 3.63, Cohen's D = 0.27, p = 0.0004, Fig. 2a) based on 7T high-resolution resting-state fMRI data from the Human Connectome Project (HCP) [21]. Further, we extracted the *t*-maps of the difference between the seed-based FC from VTA and SNc (i.e. 'VTA > SNc') (Extended Data Fig. 2a-c), which was exhibited high similarities, although in opposite directions, with the *t*-maps of both evaluation (r = 0.22, p_{adj} < 10E-20) and readiness (r = -0.12, p_{adj} < 10E-12, Fig. 2b) models. Thus, the separate VTA and SNc dopamine projections could be the putative source of evaluation and readiness signals, respectively.

Independence of decomposed neural signals

We further demonstrated that the distinct underlying neural pathways of decomposed cognitive processes were indeed independent, which could be directly inferred from uncorrelated signal components in the proposed decoding procedure (see Methods for detailed proof). Based on our simulation results, only if the compound signals were indeed a combination of independent signals, the decomposed signals from DeCoP could be uncorrelated ($|r_{mean}| < 0.001$, the 'Independent' model in Extended Data Fig. 1b & Table 2). Otherwise, the decomposed signals were highly correlated and hence inseparable (the 'One Signal' model: r_{mean} = 0.54; the 'Push and Pull' model: r_{mean} = -0.45; Extended Data Fig. 1b & Table 2). Thus, DeCoP also serves as an analytical tool for the statistical inference on the independence of decomposed signals. In this study, we found that the signals attributed to evaluation and N-shape models

(dependent signals: r_{mean} = -0.093, $p_{bootstrap}$ < 0.0001 based on 10000 bootstrap; 32.06% voxels with r < -0.1) together described the sensitivity of evaluation from punishment to reward. This was independent ($|r_{mean}|$< 0.01, $p_{bootstrap}$ > 0.3; > 99% voxels with $r \in$ [-0.1,0.1]) of those attributed to the readiness and W-shape models (dependent signals: r_{mean} = -0.159, $p_{bootstrap}$ < 0.0001; 93.29% voxels with r < -0.1) that together described the differentiated engagement of readiness from the neutral condition to reward/punishment conditions (Fig. 2a&c; see Supplementary Information for more details). Hence, the observed unbalanced sensitivity towards reward and punishment could be parsed into two independent and balanced signal components, i.e. a point-symmetric evaluation process and a line-symmetric readiness process. Moreover, the paradoxical 'inactive' vmPFC during the large-win vs neutral contrast could now be understood as a product of a trade-off between two independent processes: activation by reward stimuli (i.e. of the evaluation process) and deactivation as part of the default mode network (i.e. of the readiness process).

Complementary signal components from DeCoP

On the other hand, the N-shape and W-shape models (Fig. 1c) account for the deviation from (or equivalently, they adjust) the shapes of the proposed evaluation and readiness models, particularly the response to small reward/punishment conditions (Fig. 3c). This modification of the additional complementary orthogonal vectors also provided an effective measurement of distance between the latent pure signal and the proposed model (see Methods and Supplementary Information for more details). Therefore, converging evidence indicated that most brain regions, including VS and vmPFC, distinguish reward from punishment signals with their relative scales. In another word, the brain simplifies the scales of the different values when processing reward/punishment information, hence processing highly abstract information only. However, the bilateral anterior insula (aINS) and dorsal anterior cingulate cortex (dACC), commonly referred to as the salience network, were only sensitive to the large reward/punishment conditions, while the small reward/punishment and neutral conditions were undifferentiated, hence most likely tracking the parametric nature of

the experimental design (i.e. [-5, -0.2, 0, 0.2, 5] Extended Data Fig. 3&4 and Extended Data Table 3; see Supplementary Information for detailed analyses). These results were consistent with the role of aINS and dACC in updating and maintaining subjective value information.[22,23]

Signal components affect task performance

We further implemented a weighted voxel co-activation network analysis (WVCNA) on the ABCD MID task to capture the most informative brain-wide signal clusters [24] (see Supplementary Information for details) and identified 55 and 194 clusters for the evaluation and readiness models, respectively (Extended Data Fig. 5-6& Supplementary Table 1). Again, we observed brain-wide low correlations between evaluation and readiness clusters ($r_{mean} = 0.01$, $p = 0.1621$, > 99% pairwise $r \in [-0.1,0.1]$; Fig. 2b), hence further supporting their neural independence. We then investigated the effects of decomposed signals on task performance using canonical correlation analysis (CCA), and found associations between variations in the neural signal and task performance across three condition categories (i.e. reward, neutral and punishment) for both evaluation and readiness: accuracy (evaluation: adjusted η^2 (adj-η^2) = 0.025, $p_{perm} < 0.001$; readiness: adj-η^2 = 0.079, $p_{perm} < 0.001$) and reaction time (evaluation: adj-η^2 = 0.020, $p_{perm} < 0.001$; readiness: adj-η^2 = 0.075, $p_{perm} < 0.001$) (Supplementary Table 2&3). For the evaluation clusters, higher accuracy and faster reactions for both reward and punishment conditions (but not the neutral condition) were mainly associated with reduced sensitivity in left ventral striatum (VS), bilateral inferior temporal-occipital junctions (iTOJ), and in both sub- and pre-genual anterior cingulate cortices (sgACC and pgACC) (presented by the first components; r < -0.04, Supplementary Table 2&3, Extended Data Fig. 7a&b). For the readiness clusters, the executive regulation related to accuracy might function differently in the presence or absence of valence signals: in reward/punishment trials (represented by the first component), higher accuracy was associated with greater engagement of task-preparation regions, i.e. including the somatosensory-motor cortices and the dorsal attention network (r > 0.04) but suppression of the salience network, i.e. the dorsal aINS and the dorsal ACC (dACC)

(r < -0.04), whereas on neutral trials (represented by the second component) both relationships were reversed, with higher accuracy now being associated with suppressed task-preparation areas (r < -0.04) and greater engagement of the salience network, i.e. the ventral aINS and dACC (r > 0.04, Supplementary Table 2&Extended Data Fig. 7c). Nevertheless, faster reactions were uniformly associated with greater engagement of task-preparation areas (r > 0.04) and lesser engagement bilaterally of the anterior putamen (aPUT), left VS, dACC and the right frontoparietal network (FPN, i.e. dorsal lateral prefrontal cortex (DLPFC) and intra-parietal sulcus (IPS)) (r < -0.04) across all three conditions (represented by the first component, Supplementary Table 3&Extended Data Fig. 7d; see Supplementary Information for more information).

Decompose emotion processes with DeCoP

To further demonstrate the broader applicability of this unified strategy for signal decomposition, we decoded another complex cognitive process, emotional face processing. Similarly to reward/punishment processing, emotional processing also involves compound signals of distinct behavioural processes such as affective valence, emotional arousal, and attention, especially in the amygdala [25]. To address this issue using the DeCoP method we analysed fMRI data for a modified emotional faces task performed in 1091 19-year-old adolescents from the IMAGEN project [26]. While the right amygdala was activated by all types of facial stimuli in the emotional faces task (Fig. 4a, $t_{1090} > 13$, Cohen's D > 0.39), there was no significant difference in activations between the angry and neutral conditions (Fig. 4b upper right), consistent with previous studies [27,28]. Given the multidimensional role of the amygdala in several of the underlying processes it is conceivable that the emotional-decoding process-specific neural responses have been obscured by other processes, such as general processing of facial features. Again, we hypothesized that two dissociable components are involved in emotional processing, i.e. affective valance, corresponding to the neural representation of the affective facial emotion information (represent as [1, 0, -1] from angry to happy, Fig. 4c) and facial arousal, which was defined as the degree of excitement or motivational reaction when a person experienced and recognized an

emotional face (represent as [1, 0, 1], Fig. 4c). By applying DeCoP, we found an independent positive reaction to affective valence and negative reaction to facial arousal (Fig. 4c, r_{mean} = -0.041, $p_{bootstrap}$ = 0.1083; 97.78% voxels with $|r| < 0.1$). Hence, the right amygdala linearly tracked emotional valence on the happy through neutral to angry dimension, while the neutral faces attracted more attention (i.e. higher activation) than emotional ones. Therefore, both processes lead to an undifferentiated contrast for angry vs neutral faces.

Conclusion

In conclusion, we have developed and evaluated a universally applicable, novel signal decomposition strategy, 'DeCoP', to disentangle behavioural processes that confound the observation of functional neuroimaging signals. Through DeCoP we demonstrated the independence of evaluation and readiness processing in the brain, putatively modulated differentially by neural circuits targeting VTA and SNc; we also demonstrated that most brain regions, including the ventral striatum, coded signals based on abstract information instead of the observed exact values, except for the salience network, i.e. pgACC/dACC and aINS. Most importantly, we demonstrated that DeCoP could help to resolve common paradoxical observations in fMRI tasks which involve multiple interferential latent behavioural or cognitive processes, for example, the unexpectedly ‘inactive’ vmPFC in the contrast of large reward vs no reward, and the lack of differential amygdala activation in the contrast of angry vs neutral faces. We expect DeCoP can be applied usefully in many other comparably ambiguous data-sets.

Main Reference

1 Buschman, T. J. & Kastner, S. From Behavior to Neural Dynamics: An Integrated Theory of Attention. *Neuron* **88**, 127-144, doi:10.1016/j.neuron.2015.09.017 (2015).

2 Peixoto, D. *et al.* Decoding and perturbing decision states in real time. *Nature* **591**, 604-609, doi:10.1038/s41586-020-03181-9 (2021).

3 Fellows, L. K. The cognitive neuroscience of human decision making: a review and conceptual framework. *Behav Cogn Neurosci Rev* **3**, 159-172, doi:10.1177/1534582304273251 (2004).

4 Kahnt, T., Park, S. Q., Haynes, J. D. & Tobler, P. N. Disentangling neural representations of value and salience in the human brain. *Proc Natl Acad Sci U S A* **111**, 5000-5005, doi:10.1073/pnas.1320189111 (2014).

5 Zhang, Z. *et al.* Distributed neural representation of saliency controlled value and category during anticipation of rewards and punishments. *Nat Commun* **8**, 1907, doi:10.1038/s41467-017-02080-4 (2017).

6 Zald, D. H. & Treadway, M. T. Reward Processing, Neuroeconomics, and Psychopathology. *Annual Review of Clinical Psychology* **13**, 471-495, doi:10.1146/annurev-clinpsy-032816-044957 (2017).

7 Kragel, P. A. & LaBar, K. S. Decoding the Nature of Emotion in the Brain. *Trends Cogn Sci* **20**, 444-455, doi:10.1016/j.tics.2016.03.011 (2016).

8 Lutz, K. & Widmer, M. What can the monetary incentive delay task tell us about the neural processing of reward and punishment? *Neuroscience and Neuroeconomics*, doi:10.2147/nan.S38864 (2014).

9 Juechems, K. *et al.* A Network for Computing Value Equilibrium in the Human Medial Prefrontal Cortex. *Neuron* **101**, 977-987 e973, doi:10.1016/j.neuron.2018.12.029 (2019).

10 Kennerley, S. W., Behrens, T. E. & Wallis, J. D. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* **14**, 1581-1589, doi:10.1038/nn.2961 (2011).

11 Lebreton, M., Jorge, S., Michel, V., Thirion, B. & Pessiglione, M. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* **64**, 431-439, doi:10.1016/j.neuron.2009.09.040 (2009).

12 Tom, S. M., Fox, C. R., Trepel, C. & Poldrack, R. A. The neural basis of loss aversion in decision-making under risk. *Science* **315**, 515-518, doi:10.1126/science.1134239 (2007).

13 Roesch, M. R. & Olson, C. R. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* **304**, 307-310, doi:10.1126/science.1093223 (2004).

14 Wunderlich, K., Rangel, A. & O'Doherty, J. P. Neural computations underlying action-based decision making in the human brain. *Proc Natl Acad Sci U S A* **106**, 17199-17204, doi:10.1073/pnas.0901077106 (2009).

15 O'Doherty, J. P. The problem with value. *Neurosci Biobehav Rev* **43**, 259-268, doi:10.1016/j.neubiorev.2014.03.027 (2014).

16 Litt, A., Plassmann, H., Shiv, B. & Rangel, A. Dissociating valuation and saliency signals during decision-making. *Cereb Cortex* **21**, 95-102, doi:10.1093/cercor/bhq065 (2011).

17 Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci* **32**, 43-54, doi:10.1016/j.dcn.2018.03.001 (2018).

18 Krebs, R. M., Heipertz, D., Schuetze, H. & Duzel, E. Novelty increases the mesolimbic functional connectivity of the substantia nigra/ventral tegmental area (SN/VTA) during reward anticipation: Evidence from high-resolution fMRI. *NeuroImage* **58**, 647-655, doi:10.1016/j.neuroimage.2011.06.038 (2011).

19 Ferenczi, E. A. *et al.* Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science* **351**, aac9698, doi:10.1126/science.aac9698 (2016).

20 Lake, J. I. *et al.* Reward anticipation and punishment anticipation are instantiated in the brain via opponent mechanisms. *Psychophysiology* **56**, e13381, doi:10.1111/psyp.13381 (2019).

21 Van Essen, D. C. *et al.* The Human Connectome Project: a data acquisition perspective. *Neuroimage* **62**, 2222-2231, doi:10.1016/j.neuroimage.2012.02.018 (2012).

22 Arulpragasam, A. R., Cooper, J. A., Nuutinen, M. R. & Treadway, M. T. Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. *Proc Natl Acad Sci U S A* **115**, E5233-E5242, doi:10.1073/pnas.1800444115 (2018).

23 Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol Rev* **108**, 624-652, doi:10.1037/0033-295x.108.3.624 (2001).

24 Mumford, J. A. *et al.* Detecting network modules in fMRI time series: A weighted network analysis approach. *Neuroimage* **52**, 1465-1476, doi:10.1016/j.neuroimage.2010.05.047 (2010).

25 Simic, G. *et al.* Understanding Emotions: Origins and Roles of the Amygdala. *Biomolecules* **11**, doi:10.3390/biom11060823 (2021).

26 Schumann, G. *et al.* The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol Psychiatry* **15**, 1128-1139,

doi:10.1038/mp.2010.4 (2010).

27 Fusar-Poli, P. *et al.* Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J Psychiatry Neurosci* **34**, 418-432 (2009).

28 Quinlan, E. B. *et al.* Psychosocial Stress and Brain Function in Adolescent Psychopathology. *Am J Psychiatry* **174**, 785-794, doi:10.1176/appi.ajp.2017.16040464 (2017).

Figures

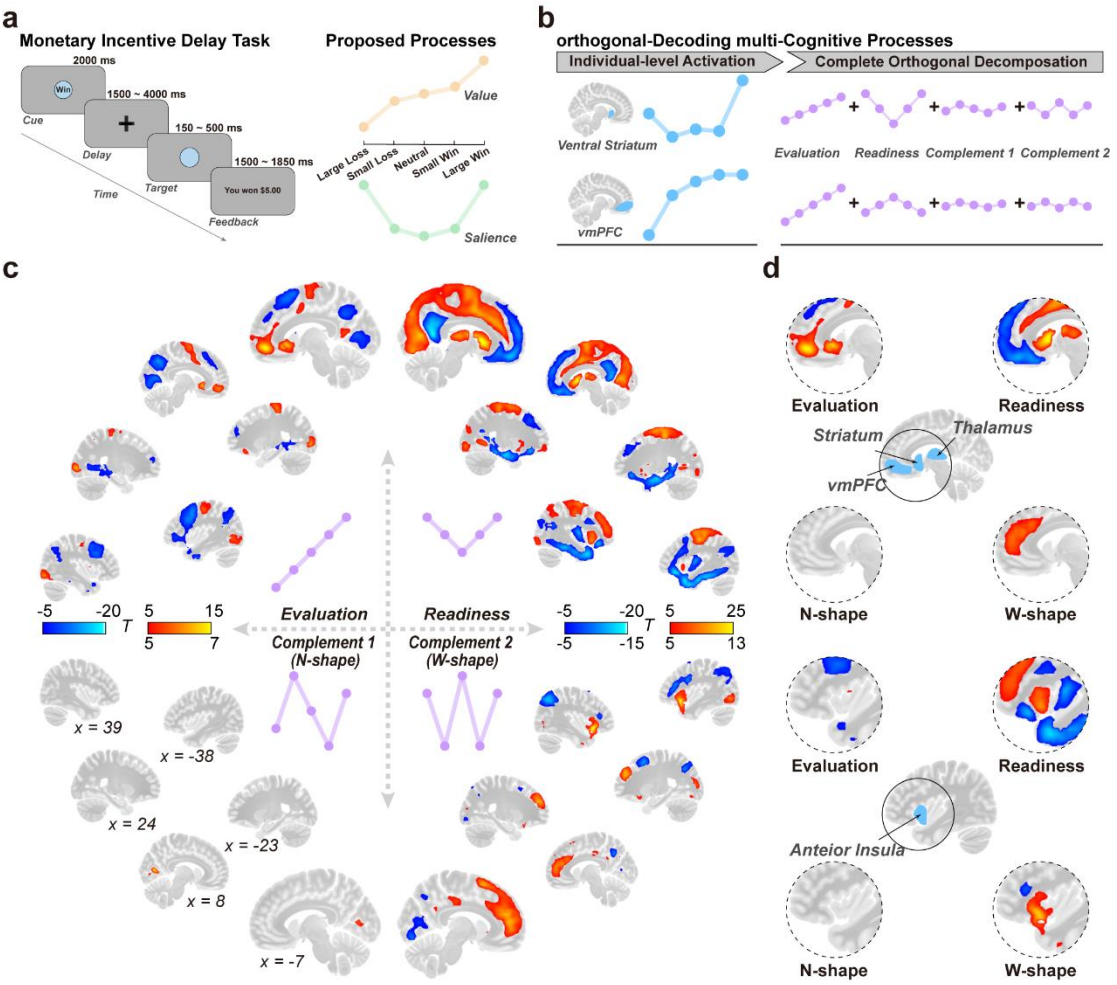


Fig. 1. Experimental design and neural representations of the orthogonal decomposition. a. Procedure of the monetary incentive delay task (MID) and the proposed evaluation and readiness cognitive processes; **b.** An illustration of Orthogonally Decoding multi-Cognitive Processes (DeCoP); **c.** Brain-wide T-maps of decomposed signals for orthogonal contrasts. Brain-wide significance was set as |T|>5. The MNI coordinates of brain slices were inserted at the lower left; d. Decomposed signals in highlighted brain regions.

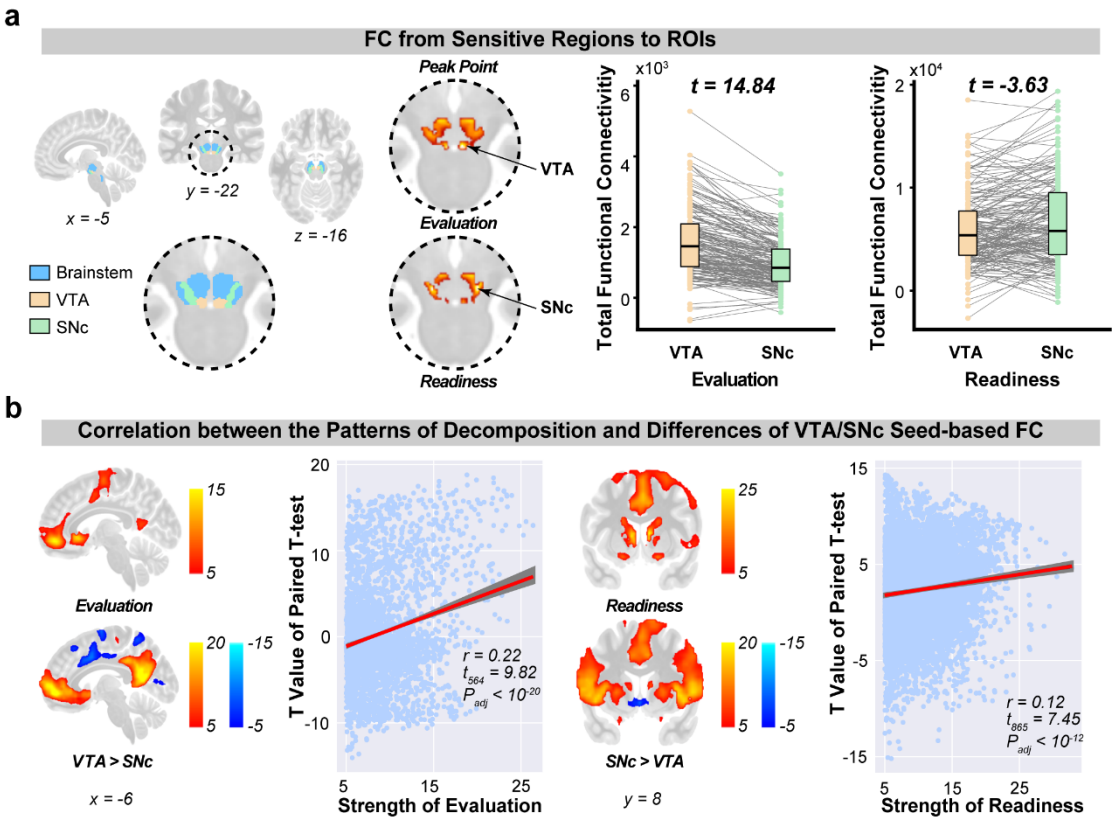


Fig. 2 The decomposed evaluation and readiness processing targets VTA and SNc neural circuits respectively. a. Left: the masks of Brainstem, VTA and SNc were from AAL3 atlas. Middle: strength of functional connectivities (FCs) to VTA and SNc from the evaluation and readiness regions identified in Fig. 1. Right: Paired t-tests between FCs to VTA and SNc from evaluation and readiness; **b.** Brain-wide pattern correlations between the strength of decomposition signals (left: evaluation; right: readiness) and the differences of seed-based FCs from VTA and SNc. Brain-wide significance was set as |T|>5.

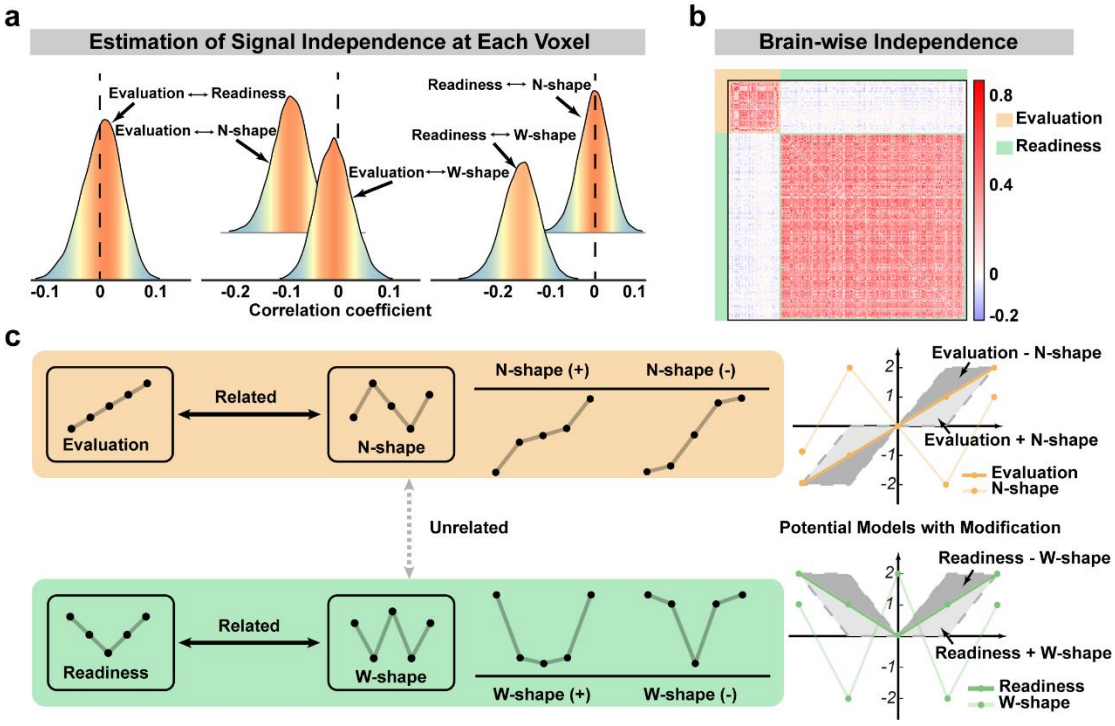


Fig. 3 Inference for independence among orthogonally decomposed signals during the reward anticipation. a. The distributions of pair-wise correlations between signals of orthogonal contrasts at each voxel. Mean correlations deviated from 0 would indicate a pair of related signals; **b.** The correlation matrix of signals from the evaluation and readiness clusters identified by WVCNA. **c.** An illustration of how related signals could describe the evaluation-related and readiness-related processing.

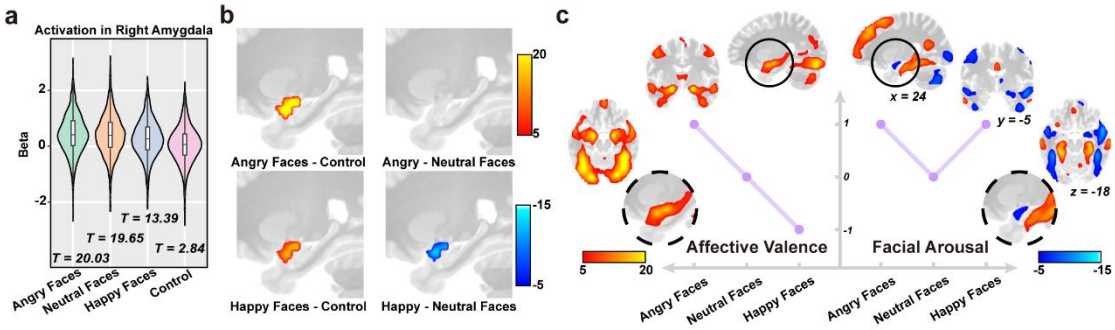


Fig. 4 Decoding affective valence and facial arousal of emotional face processing in amygdala by DeCoP. a. BOLD signals of the right amygdala under different task conditions. **b.** Activations of the right amygdala under different contrasts. **c.** The neural representations of affective valence and facial arousal based on orthogonal decompositions. Brain-wide significance was set as |T|>5.

Methods

Participants

The dataset used for this study was selected from Annual Curated Data Release 2.01 (https://data-archive.nimh.nih.gov/abcd) of the Adolescent Brain Cognitive Development (ABCD) cohort, which recruited 11,875 children between 9–11 years of age from 21 sites across the United States [17]. The study conforms to each site's Institutional Review Board's rules and procedures, and all participants provide informed consent (parents) or informed assent (children). More details of the subjects and the data collection are provided at the ABCD website (https://abcdstudy.org/scientists/protocols) and are also described previously [17].

Magnetic resonance imaging (MRI) data in the ABCD study were collected from different 3T scanner platforms (i.e. Siemens Prisma, General Electric (GE) MR750 and Philips Achieva dStream). To minimize biases introduced by multiple platforms, we only included MRI data from the most frequent manufacturer Siemens Prisma, i.e. 5968 participants from 13 sites. By examining the similarity of brain activations across these 13 sites, we selected 2326 participants from 4 sites with consistent activation patterns. Furthermore, the data with poor registration (by visual check) and high head motion (mean framework displacement (FD) > 0.5 mm) were excluded. Hence, 1939 quality-controlled participants were included in the following analysis and the demographic characteristics of these participants are summarized in Extended Data Table 1.

Monetary incentive delay (MID) task design

A modified version of the MID task was used to examine brain activation during monetary reward anticipation and receipt [29], which consists of five levels of incentive: large loss, small loss, neutral, small win and large win (i.e., -5.0 $, -0.2 $, 0$, 0.2 $ and 5.0$ respectively). In each trial, participants were first presented with one of three cue shapes (circle, square or triangle) that indicated the trial condition (of win, loss, or neutral, respectively), as well as the amount of money involved. This cue presentation (2,000 ms) was followed by a jittered anticipatory delay (1,500-4,000 ms) of fixation on a black crosshair. Subsequently, a blank target cue (with the same shape as the

previously presented cue) emerged and required the participants to press a response button before the target disappeared to win or avoid loss. With a tracking algorithm, the time of the target on the screen was dynamically manipulated (i.e., 150 ms-500 ms) to maintain a 60% success rate for each participant. After a short delay, feedback of the current trial (i.e., the amount of monetary gain or loss) and the accumulated reward so far were presented for 1,500-1,850 ms (Fig. 1A). During the anticipation phase, participants underwent 50 trails in total (i.e., 10 trails per incentive degree). Participants had first completed a practice session outside the scanner before completing two sessions of the MID task with fMRI recording (approximately 5.5 minutes each).

Image acquisition and preprocessing

Imaging protocols were harmonized across sites and scanners. 3-dimensional T1-weighted images (1.0 mm isotropic, TR = 2500 ms, TE = 2.88 ms) were acquired with a gradient-echo sequence for anatomical localization and high spatial (2.4 mm isotropic) and temporal (TR = 800 ms) resolution MID-task-based fMRI was acquired with echo-planar imaging (EPI) sequence in two separate runs (approximately 5.5 minutes each). The detailed MRI acquisition protocol is described elsewhere [17]. All functional images were preprocessed with the same preprocessing procedure by suggested protocols from FMRIB's Software Library (FSL v5.0.9), Advanced Normalization Tools (ANTs v1.9.2) and Analysis of Functional NeuroImages (AFNI v18.3.03). Concretely, the whole preprocessing procedure included the following steps: (i) brain extraction (ANTs antsBrainExtraction), nonlinear registration to MNI space (ANTs antsRegistrationSyN) for structural images; (ii) rigid realignment to adjust for motion (FSL MCFLIRT) and field map correction (FSL TOPUP) for functional images; (iii) co-registration to a high-resolution T1 image and normalization to 3 mm isotropic MNI standard space (ANTs antsRegistrationSyN) and (iv) spatial smoothing with a 6mm full-width at half-maximum (FWHM) Gaussian kernel (AFNI 3dBlurToFWHM) and detrending (AFNI 3dDetrend).

First-level analysis of task-based fMRI

At the first-level analysis, we set up a general linear model to estimate the effects of the task conditions at the individual level with SPM, which contained 20 task condition regressors (i.e. Target (Hit or Miss) * Phases (Anticipation, Feedback) * Task Conditions (Large-Loss, Small-Loss, Neutral, Small-Win or Large-Win)) and additional covariate regressors (i.e., 24 motion-related parameters: 6 rigid-body motion parameters, their first temporal derivatives and 12 corresponding squared items; and mean signals of both white matter and ventricles). All regressors were convolved with a double-gamma HRF, and no orthogonalization of regressors was applied. For BOLD signals, drift was modeled with the DCT basis, and a cut-off of 128 s was applied (SPM defaults). Finally, the autocorrelation was modeled as a global approximate AR(1) in SPM.

Orthogonally Decoding multi-Cognitive Processes

In the present study, we propose a novel approach to decompose each participant's brain activations at varied conditions (denoted as y) with a set of orthogonal basis $\mathbf{x}=(x_i,\dots x_k)$, where each vector could represent a predefined signal model, e.g. evaluation or readiness. Specifically, 'orthogonal' here means that the pairwise covariances of vectors were all zero, i.e. $Cov(x_i,x_j)=E(x_i,x_j)-E(x_i)E(x_j)=0$. In this way, the regression coefficients $\boldsymbol{\beta}=(\beta_1,\dots\beta_k)$ (i.e. the strength of signals for each individual) estimated from a multiple linear model with all vectors were the same as those estimated univariately (of simple linear models), i.e. $T(\beta_i \mid y,\mathbf{x},\boldsymbol{\beta}_{-i})=T(\beta_i \mid y,x_i)$, where $T(\cdot)$ stands for the best linear unbiased estimator.

We propose that the above individual level orthogonal disconfounding eliminates spurious correlations of signal components (i.e. β) introduced by related vectors (i.e. X_i are correlated), thus allowing us to make meaningful inferences regarding signal independence at the population level. Here, we only describe a simple proof for the purpose of illustration, where only two orthogonal vectors were involved. Let y_i denote the activations of the ith individual across all conditions, and x_{1i} and x_{2i} denote the two

predefined independent orthogonal models of input signals, where $Cov(x_{1i}, x_{2i}) = 0$, and β_{1i} and β_{2i} denotes the corresponding regression coefficients, i.e. the signal strengths. We have the following linear model:

$$y_i = \beta_{1i} x_{1i} + \beta_{2i} x_{2i} + \varepsilon . \quad (1)$$

We first show that orthogonal settings of x_{1i} and x_{2i} are necessary for a meaningful signal decomposition. The covariance and variance conditional on x_{2i} are easily derived.

$$Cov(x_{1i}, y_i \mid x_{2i}) = Cov(x_{1i}, y_i) - Cov(x_{1i}, x_{2i}) \cdot Cov(y_i, x_{2i}) \Big/ Var(x_{2i})$$

and

$$Var(x_{1i} \mid x_{2i}) = Var(x_{1i}) - Cov(x_{1i}, x_{2i})^2 \Big/ Var(x_{2i}) .$$

Similar results are easily acquired for the covariance/variance conditional on x_{1i}. Therefore, we establish the following relationship between the least square estimations (LSEs) of β_{1i} and β_{2i}:

$$\begin{aligned}
&\frac{Cov(x_{1i}, y_i)}{Var(x_{1i})} - \beta_{1i} \\
&= \frac{Cov(x_{1i}, y_i)}{Var(x_{1i})} - \frac{Cov(x_{1i}, y_i \mid x_{2i})}{Var(x_{1i} \mid x_{2i})} \\
&= \frac{Cov(x_{1i}, y_i)}{Var(x_{1i})} - \frac{Cov(x_{1i}, y_i) - \frac{Cov(x_{1i}, x_{2i}) \cdot Cov(x_{2i}, y_i)}{Var(x_{2i})}}{Var(x_{1i}) - \frac{Cov(x_{1i}, x_{2i})^2}{Var(x_{2i})}} \\
&= \frac{Cov(x_{1i}, x_{2i}) \cdot Cov(x_{2i}, y_i) \cdot Var(x_{1i}) - Cov(x_{1i}, y_i) \cdot Cov(x_{1i}, x_{2i})^2}{Var(x_{1i})^2 \cdot Var(x_{2i}) - Var(x_{1i}) \cdot Cov(x_{1i}, x_{2i})^2} \\
&= \frac{Cov(x_{1i}, x_{2i})}{Var(x_{1i})} \cdot \frac{Cov(x_{2i}, y_i) - \frac{Cov(x_{1i}, x_{2i}) \cdot Cov(x_{1i}, y_i)}{Var(x_{1i})}}{Var(x_{2i}) - \frac{Cov(x_{1i}, x_{2i})^2}{Var(x_{1i})}} \\
&= \frac{Cov(x_{1i}, x_{2i})}{Var(x_{1i})} \cdot \frac{Cov(x_{2i}, y_i \mid x_{1i})}{Var(x_{2i} \mid x_{1i})} \\
&= \frac{Cov(x_{1i}, x_{2i})}{Var(x_{1i})} \cdot \beta_{2i}
\end{aligned} .$$

Clearly, unless $Cov(x_{1i}, x_{2i}) = 0$, β_{1i} could always be expressed as a function of β_{2i}, and hence can never be independent of each other. The above derivations, therefore, prove the necessity of using orthogonal vectors for meaningful signal decomposition.

Meanwhile, by setting $Cov(x_{1i}, x_{2i}) = 0$, and realizing that the LSE of β_{1i} in the absence of x_{2i} from the regression model (1) (i.e. reduced to a simple linear model) is $\beta'_{1i} = Cov(x_{1i}, y_i)/Var(x_{1i})$, we could immediately have $\beta'_{1i} = \beta_{1i}$. Thus with orthogonal x_{1i} and x_{2i}, estimations from a multiple linear model would be the same as those univariately estimated from simple linear models.

We then show that the population-level correlation analyses of β_{1i} and β_{2i} derived above provide meaningful statistical inferences of signal independence. For simplicity, assume that we rewrite β_{2i}, the signal strength of model x_{2i}, into a sum of two independent components, i.e., $\beta_{2i} = \beta'_{2i} + \gamma_i \beta_{1i}$ where β'_{2i} is independent of β_{1i}, i.e., $Cov(\beta'_2, \beta_1) = 0$, and the parameter γ_i denotes the proportion of overlapped signals with model x_{1i}, and is independent of β_{1i}, i.e., $Cov(\gamma, \beta_1) = 0$. We then calculate the population-level correlation of β_{1i} and β_{2i} as:

$$
\begin{aligned}
Cov(\beta_1, \beta_2) &= Cov(\beta_1, \beta'_2 + \gamma\beta_1) \\
&= Cov(\beta_1, \beta'_2) + Cov(\beta_1, \gamma\beta_1) \\
&= E(\gamma\beta_1^2) - E(\beta_1)E(\gamma\beta_1) \\
&= E(\gamma\beta_1^2) - E(\gamma)E(\beta_1^2) + \left(E(\gamma)E(\beta_1^2) - E(\beta_1)E(\gamma)E(\beta_1)\right), \\
&= Cov(\gamma, \beta_1^2) + E(\gamma)Var(\beta_1) \\
&= E(\gamma)Var(\beta_1)
\end{aligned}
$$

which only equals 0 if either $E(\gamma) = 0$, i.e. the population mean signals of both models are not a function of each other, or $Var(\beta_1) = 0$, i.e. one of the signals is invariant across the population, and hence again is not dependent on each other. Either way, the signals of both models are indeed independent, and thus $Cov(\beta_1, \beta_2) = 0 \Leftrightarrow \beta_1 \perp \beta_2$, where orthogonal decomposition at the individual level permits the expression

$Cov(\beta_1, \beta_2) = 0$. In conclusion, this two-stage approach can be used for statistical inference concerning the signal dependence or independence of two signals.

To illustrate the above theoretical derivations, we conducted simulation analyses to evaluate three models, namely the "independent-signal model", where each observed signal consists of two independent inputs ($\beta_{1i} \perp \beta_{2i}$, and hence $Cov(\beta_1, \beta_2) = 0$), and two dependent-signal models, i.e. the 'single signal model' with $\beta_{2i} = \gamma_i \beta_{1i}$ (assume $E(\gamma) > 0$, and hence $Cov(\beta_1, \beta_2) > 0$), and the 'push-and-pull model' with $\beta_{2i} + \gamma_i \beta_{1i} = c_i$ (assume $E(\gamma) > 0$, and hence $Cov(\beta_1, \beta_2) < 0$). For simplicity, the signal overlapping parameter γ_i was set as a constant (i.e. 0 for the independent-signal model and 1 in both dependent-signal models); the observed signals (y in equation 1) were simulated from [-2, -1, 0, 1, 2] to [2, 1, 0, 1, 2] with an increment of [0.2, 0.1, 0, 0, 0] at each step, i.e. [-2, -1, 0, 1, 2], [-1.8, -0.9, 0, 1, 2], …, [1.8, 0.9, 0, 1, 2], [2, 1, 0, 1, 2]; and the orthogonal vectors were fixed as x_{1i} = [2, 1, 0, 1, 2] and x_{2i} = [-2, -1, 0, 1, 2]. Therefore, for the 'single signal model', y_i was directly set as the observed signals plus a random noise $N(0,1)$. For the 'push-and-pull model', at the *i*th simulation step, by setting $c_i = 0.05 \times (i-1) + N(0,1)$, $\beta_{1i} = c_i + N(0,1)$ and $\beta_{2i} = 1 - c_i + N(0,1)$, we could then simulate $y_i = \beta_{1i} x_{1i} + \beta_{2i} x_{2i}$. Clearly, at the population-level, we have $Cov(\beta_1, \beta_2) = E_i^2(c) - E_i(c^2) = -Var(c)$, and hence β_1 and β_2 are dependent. For the 'independent-signal model', at the *i*th simulation, by setting $c = 0.05 \times (i-1)$, $\beta_{1i} = c + N(0,1)$ and $\beta_{2i} = 1 - c + N(0,1)$, we could also simulate $y_i = \beta_{1i} x_{1i} + \beta_{2i} x_{2i}$, where, however, one would expect $Cov(\beta_1, \beta_2) = 0$. It is notable that while the simulation models of 'push-and-pull' and 'independent-signal' are rather similar, they are fundamentally different. The reason lies in the fact that the constant c in the 'independent-signal model' is invariant across individuals and solely determined by the predefined form of observed signals, and hence the signal strengths β_{1i} and β_{2i} do not dependent on each other. However, in the 'push-and-pull model', the form of

observed signals only determined the expectation $E_i(c)$, and hence both β_{1i} and β_{2i} are nevertheless the functions of c_i that varies across individuals, and thus are not independent.

For each model, we simulated with 1000 independent individuals for 1000 times, and the detailed results of the simulation are shown in Extended Data Fig. 1b and Extended Data Table 2.

Functional connectivity (FC) based on resting-state fMRI data

A total of 184 participants' preprocessed high-resolution (7T) resting-state fMRI data were collected from the Human Connectome Project (HCP) dataset [21]. Total FC (Pearson Correlation) from evaluation and readiness activation activations regions to the voxels in brainstem (Fig. 4A), which were masked with AAL3 atlas [30] and from VTA and SNc to the whole brain were examined respectively. A paired t-test was used to estimate which brain region had a stronger FC with the activation region. The similarity between patterns of activation and FC was represented by Pearson's correlation coefficient r. Considering the high correlation between voxels, the degree of freedom was adjusted according to the number of components which had over 95% interpretation of brain activation by principal component analysis.

Measurement of modifying effects

As N-shape and W-shape models can be considered as modifiers of evaluation and readiness models, respectively, their strength (i.e. the activation level) can be used to measure the deviance from the proposed model settings, i.e. the exact scale or the relative scale. Thus we computed the standardized mean $((t_{relative} + t_{exact})/\max(abs(t_{relative}), abs(t_{exact})))$ to evaluate the prevailing settings, where $t_{relative}$ represents the t-statistics of the activation for N-shape or W-shape models under the relative-scale setting, and t_{exact} represents the t-statistics of the activation for N-shape or W-shape models under the relative-scale setting. When the observed signal falls somewhere between the relative-scale and exact-scale settings, it is readily observable that $t_{relative}$ would be positive and t_{exact} would be negative. Therefore, $abs(t_{relative}) > abs(t_{exact})$ would prefer the exact-scale setting, and the

standardized mean ranges from 0 to 1, whereas $abs(t_{relative}) < abs(t_{exact})$ would prefer the relative-scale setting, and the standardized mean ranges from -1 to 0. It is notable that meaningful settings other than this may also be possible. For instance, the response of the small stimuli could be identical to the large stimuli in the observed signal, thus $t_{exact} < t_{relative} < 0$, and the standardized mean could be well smaller than -1.

Weighted voxel co-activation network analysis (WVCNA)

The R package WGCNA [31] was implemented to conduct the WVCNA [32], which identifies activation modules of both evaluation and readiness models across the brain. The final dataset used for WVCNA included 1,912 participants with 3365 voxels for the evaluation model and 1900 participants with 10932 voxels for the readiness model after removing null data and outliers. We transferred most parameters as default settings from previous studies [33], except for the soft-threshold parameters, which were set to seven based on the scale-free topology criteria (Fig. S5), incidentally identical to those estimated for the MID task from a different cohort [33]. The stabilities of the generated modules were assessed through bootstrapping.

Canonical correlation analysis (CCA)

CCA has been widely used to investigate the overall correlation between two sets of standardized variables [34]. Due to high intra-correlations in both brain networks and task performances, multicollinearity was a potential risk factor jeopardizing the validity of the subsequent statistical inference. Therefore, we adopted the CCA proposed in a previous reference [33]. Briefly, we used the eta square (η^2) to represent the proportion of mutually explained variance between the two sets of variables. For each correlation, the *P* value or significance level was determined using the permutation test, where the individual IDs of task performances were randomly shuffled at each iteration to generate the null distribution of the corresponding test statistics. We further included an adjusted η^2 to correct for the inflation in η^2 caused by the increased number of variables as:

$$\eta^2_{adj} = 1 - \frac{1-\eta^2}{1-\eta^2_0}$$

where η_0^2 represents the expected η^2 under the null hypothesis that there is no relationship between the two sets of variables (that is, it acts as a measure of inflation in η^2), and can be directly estimated through the permutation test. As the effect of one task performance was significant, we also estimated the effects of each component of the behavior with the significant levels at $P < 0.05$.

Emotional faces task in the IMAGEN project

Emotional face processing was investigated using the "emotional faces task" in the IMAGEN project [26]. Participants were exposed to a sequence of stimuli which consisted of short (2–5 s) black-and-white video clips showing male and female faces with varying facial expressions. Stimuli showed human faces which started with the expression of a neutral expression and then either turned angry/happy or displayed a neutral movement without a particular emotional content (for example, twitching the nose). Stimuli were arranged in 18 s blocks, each block including 4–7 video clips depicting faces of the same emotion or neutral. Altogether, there were 3 blocks of neutral faces and 3 blocks containing angry and happy faces respectively. In between 2 blocks of face clips, an 18 s non-biological control video clip was presented. The control stimuli consisted of expanding and contracting black-and-white concentric circles of various contrasts, roughly matching the contrast and motion characteristics of the faces clips.

The recruitment procedures employed in the IMAGEN project and demographic information have been described previously [26]. The standard operating procedures for the IMAGEN project are available at http://www.imagen-europe.com/en/Publications_and_SOP.php, which contain details on ethics, recruitment, neuropsychological tests and preprocessing protocols of MRI data. In brief, task-based functional MRI data were analyzed with SPM12 (Statistical Parametric Mapping, http://www.fil.ion.ucl.ac.uk/spm). Preprocessing included: slice time correction to adjust for time differences due to multi-slice imaging acquisition, realignment to the first volume in line, non-linearly warping to the MNI space (based on a custom EPI template (53x63x46 voxels) created out of an average of the mean

images of 400 adolescents), resampling at a resolution of 3x3x3mm^3 and smoothing with an isotropic Gaussian kernel of 5 mm full-width at half-maximum. And the activations of each task condition, i.e. happy, angry, neutral faces and control stimuli, were evaluated for each individual at the first-level analysis (The details can be found at https://github.com/imagen2/imagen_processing). And these results were conducted for the further DeCoP analysis.

Data availability

ABCD data are available from a dedicated database: https://abcdstudy.org. Human Connectome Projects data are available from: https://www.humanconnectome.org. The IMAGEN project are available from a dedicated database: https://imagen2.cea.fr.

Code availability

Custom code that supports the findings of this study is available from the corresponding author upon request. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Information. Additional data related to this paper may be requested from the authors.

Methods Reference

29 Knutson, B., Westdorp, A., Kaiser, E. & Hommer, D. FMRI Visualization of Brain Activity during a Monetary Incentive Delay Task. *NeuroImage* **12**, 20-27, doi:10.1006/nimg.2000.0593 (2000).

30 Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J. & Joliot, M. Automated anatomical labelling atlas 3. *Neuroimage* **206**, 116189, doi:10.1016/j.neuroimage.2019.116189 (2020).

31 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).

32 Mumford, J. A. *et al.* Detecting network modules in fMRI time series: a weighted network analysis approach. *Neuroimage* **52**, 1465-1476, doi:10.1016/j.neuroimage.2010.05.047 (2010).

33 Jia, T. *et al.* Neurobehavioural characterisation and stratification of reinforcement-related behaviour. *Nat Hum Behav* **4**, 544-558, doi:10.1038/s41562-020-0846-5 (2020).

34 Combes, S. *et al.* Relationships between sensory and physicochemical measurements in meat of rabbit from three different breeding systems using canonical correlation analysis. *Meat Sci* **80**, 835-841, doi:10.1016/j.meatsci.2008.03.033 (2008).

Acknowledgments

This work received support from the following sources: National Key Research and Development Program of China (No. 2019YFA0709502 and No. 2018YFC1312900), the National Natural Science Foundation of China (No. 91630314 and No 81801773), the 111 Project (B18015), the Key Project of Shanghai Science &Technology Innovation Plan(16JC1420402), Shanghai Municipal Science and Technology Major Project (2018SHZDZX01) and Zhangjiang Lab and the Shanghai Pujiang Project (No. 18PJ1400900). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author Contributions:

Conception or Design of the Study: T.J., T.W.R. and J.F.. Manuscript Writing and Editing: S.X. and T.J. wrote the manuscript; T.W.R. and J.F. edited the first draft; all authors critically reviewed the manuscript. Imaging Data Preprocessing: S.X., C.X. and W.C.. Visualization: S.X., C.X. and T.J.. Data Analysis: S.X., C.X., Z.Z., J.K. and G.S. conducted all the statistical analyses, under the instruction of T.J. and J.F.. Results Interpretation: T.J., T.W.R. and J.F.. Supervision of the Study: T.J. and J.F.. Funding Acquisition: T.J. and J.F..

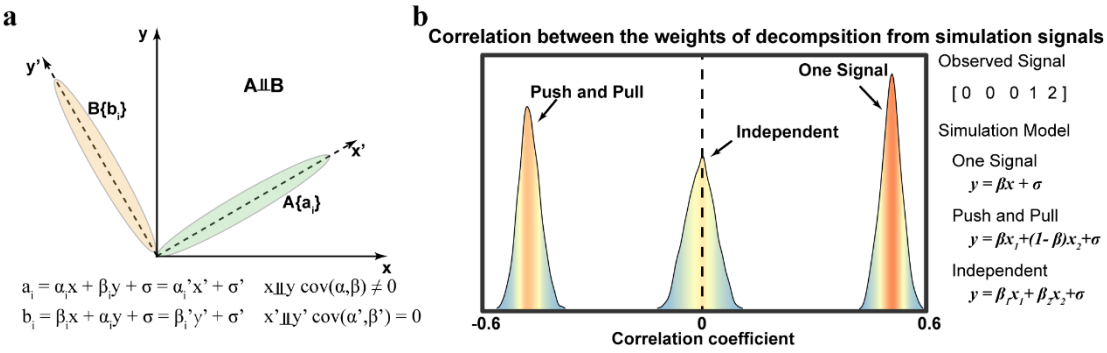
Competing Interests:

The authors declare no competing interests.

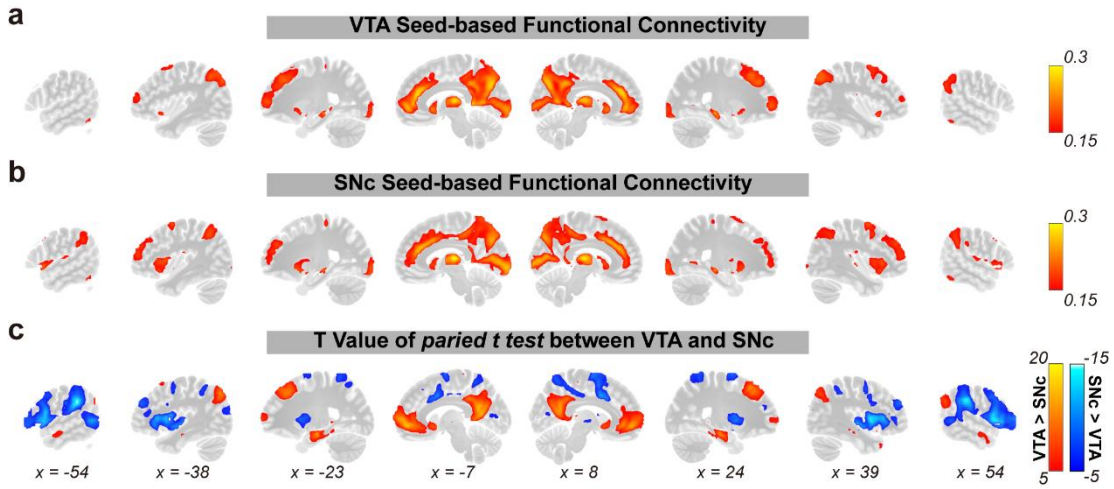
Additional Information

Supplementary Information is available for this paper.

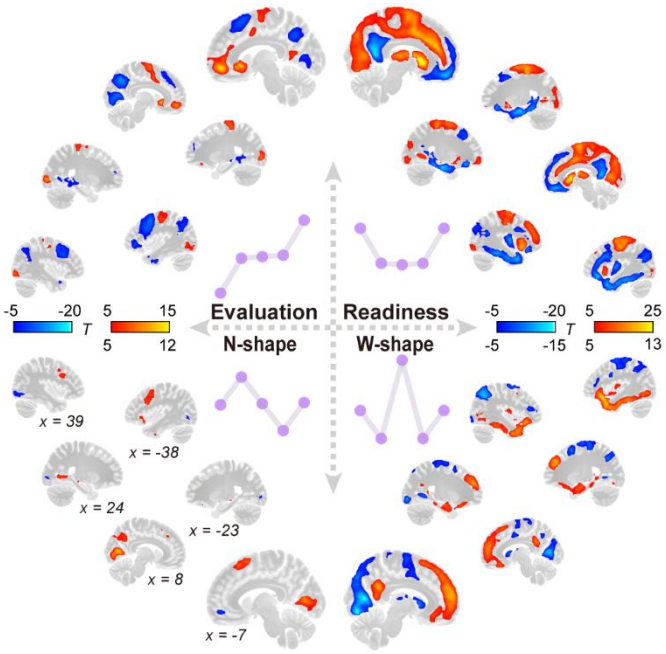
Extended Data figures and tables



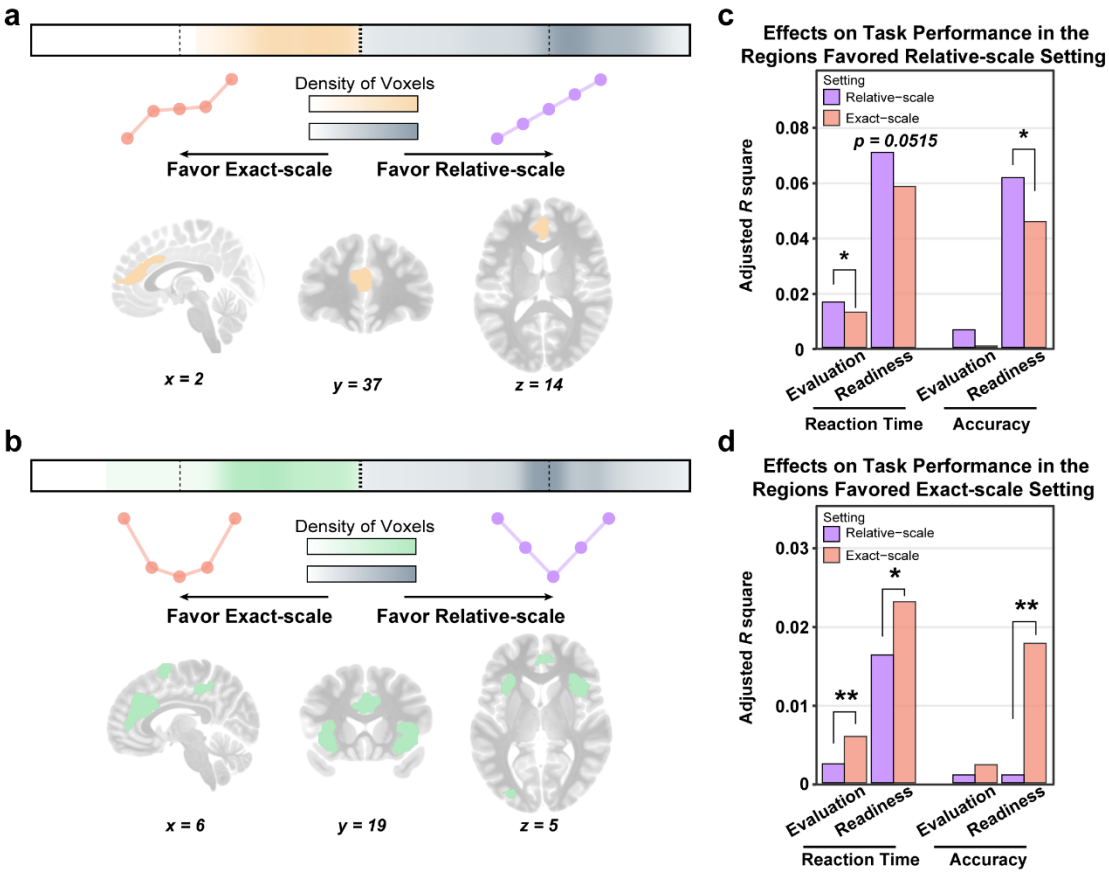
Extended Data Fig. 1. a. Only the "correct" orthogonal vectors could retrieve latent independent signals. **d.** Correlations between decomposed signals based on different simulation models. Also see Extended Data Table 2.



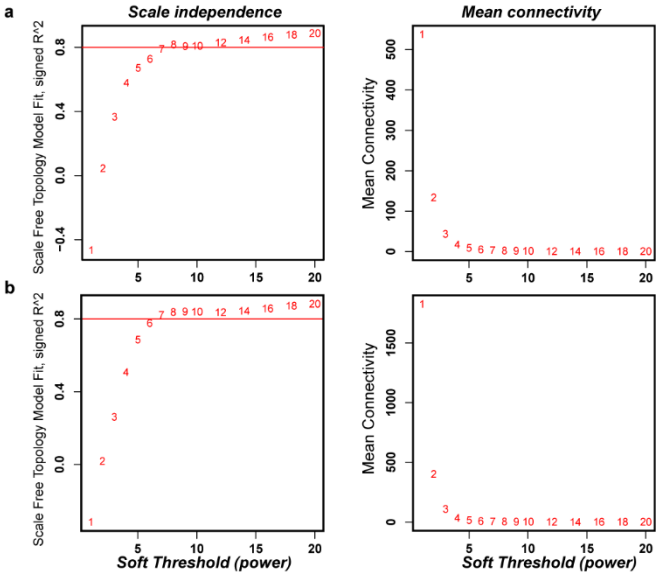
Extended Data Fig. 2 The brain-wide patterns of seed-based functional connectivity (FC) strength from regions of interest (ROIs): **a.** VTA seed-based; **b.** SNc seed-based; **c.** The T-map of the differences between seed-based FCs patterns from VTA and SNc.



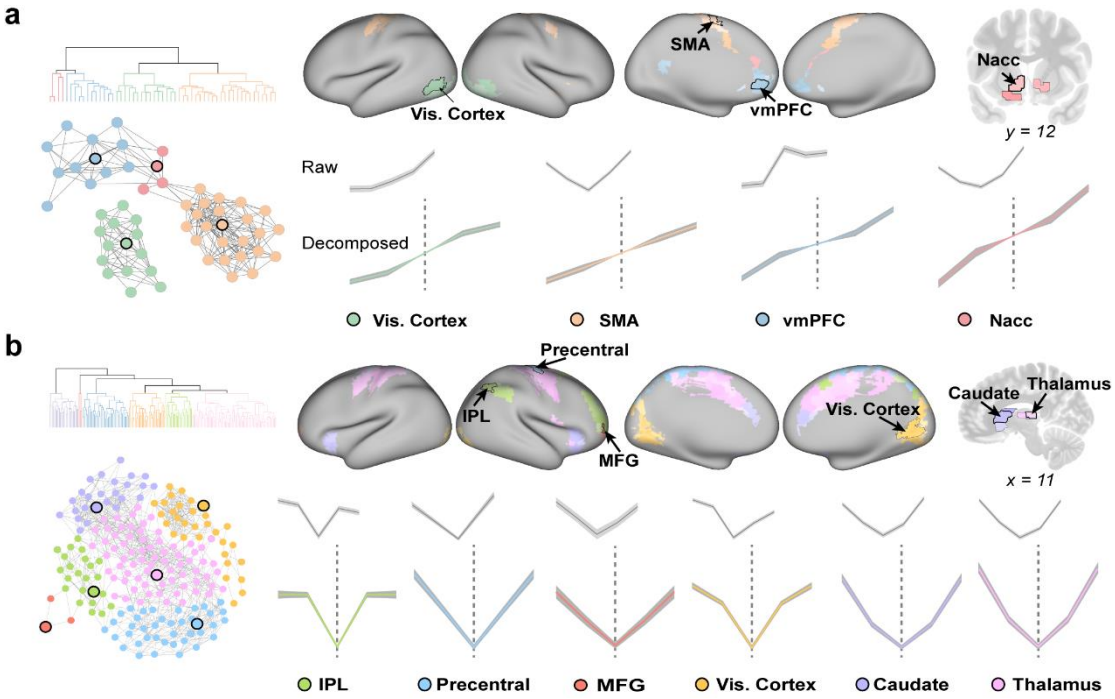
Extended Data Fig. 3 Brain-wide neural representations of signal decomposition with the exact-scale setting (e.g. the evaluation contrast takes the form [-5, -0.2, 0, 0.2, 5] from large-loss to large-win, consistent with the parametric nature of the exact monetary magnitude in the experimental design).



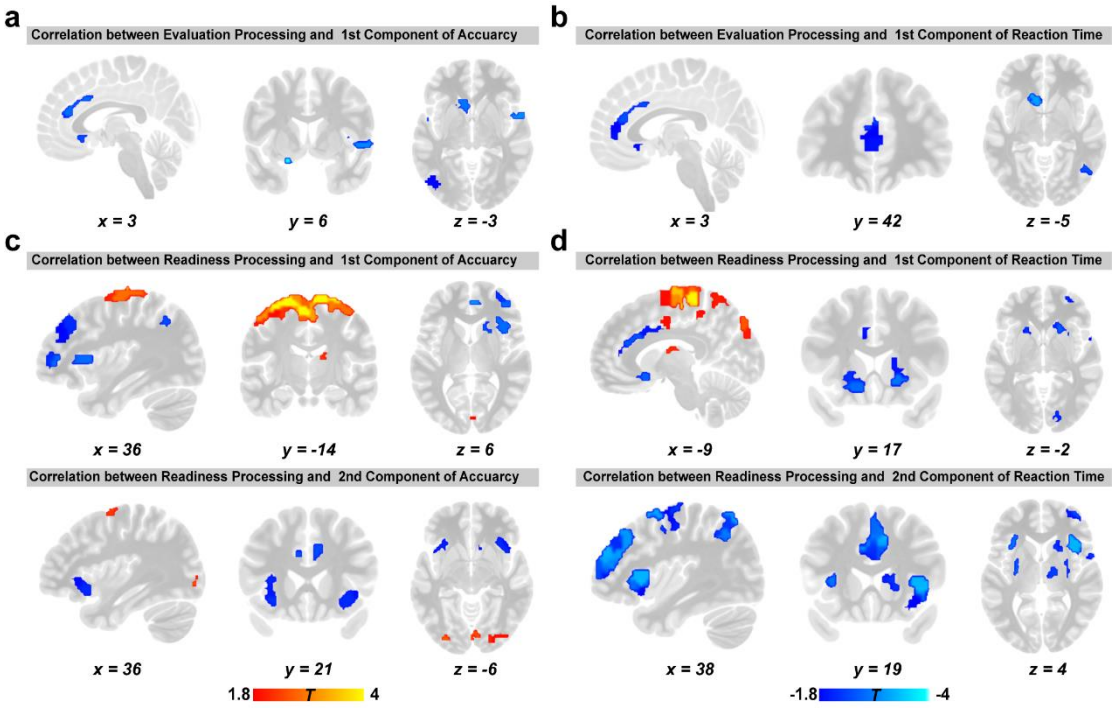
Extended Data Fig. 4 The bimodal distribution of voxels favoring the exact-scale or relative-scale settings in evaluation **(a)** and readiness **(b)** processing. The typical regions favoring the exact-scale setting were illustrated in the corresponding lower subplots. **c&d.** The favored neural representations (i.e. relative-scale vs exact-scale) demonstrated better predictions for task performance. *significant at level 0.05, ** significant at level 0.01.



Extended Data Fig.5 Plots of soft-threshold for WVCNA.



Extended Data Fig. 6 The brain clusters and subnetworks of evaluation **(a)** and readiness **(d)** processing. Also see Extended Data Table 3. The core brain region of each subnetwork is marked by a black bound and illustrated in both the networks (left) and brain templates (right). Raw and decomposed signals of the core brain regions are presented below the brain.



Extended Data Fig. 7 The T-maps of evaluation and readiness processing's impacts on the task performances. The canonical correlation analysis (CCA) was implemented to further segregate signal components that demonstrated differential associations with task performances under different experimental conditions.

703 **Extended Data Table 1** Demographic characteristics of the samples in this study.

Characteristics	Total (n=1939)	Site1 (n= 415)	Site2 (n= 362)	Site3 (n= 396)	Site4 (n= 766)
Gender (male)	984 (50.75%)	219 (52.77%)	169 (46.69%)	191 (48.23%)	405 (52.87%)
Race/ethnicity					
White	1473 (75.97%)	282 (67.95%)***	246 (67.96%)**	313 (79.04%)	632 (82.51%)***
Black	39 (2.01%)	3 (0.72%)	10 (2.76%)	18 (4.55%)**	8 (1.04%)
Hispanic	233 (12.02%)	84 (20.24%)***	52 (14.36%)	24 (6.06%)***	73 (9.53%)
Asian	20 (1.03%)	2 (0.48%)	13 (3.59%)***	4 (1.01%)	1 (0.13%)*
Other/Missing	174 (8.97%)	44 (10.6%)	41 (11.33%)	37 (9.34%)	52 (6.79%)

Superscripted * means the variable differ significantly from the total population on the corresponding variable. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by Fisher's Exact Tests.

704

Extended Data Table 2 Simulation results of the correlations between the weights of decomposition.

Observerd Signals	One Signal Model			Push and Pull Model			Independent Model		
	r	*T* (Readiness)	*T* (Evaluation)	*r*	*T* (Readiness)	*T* (Evaluation)	*r*	*T* (Readiness)	*T* (Evaluation)
2 1 0 1 2	-0.001	30.232	0.032	-0.448	21.868	0.024	0.010	30.185	0.025
1.8 0.9 0 1 2	0.080	30.033	2.669	-0.449	20.768	1.033	0.000	28.696	1.353
1.6 0.8 0 1 2	0.155	29.875	5.245	-0.449	19.667	2.086	0.000	27.154	2.751
1.4 0.7 0 1 2	0.230	29.652	7.735	-0.449	18.620	3.113	-0.002	25.624	4.087
1.2 0.6 0 1 2	0.296	29.433	10.073	-0.449	17.437	4.138	0.000	24.137	5.484
1 0.5 0 1 2	0.355	29.158	12.230	-0.450	16.356	5.174	-0.001	22.665	6.785
0.8 0.4 0 1 2	0.410	28.779	14.196	-0.449	15.333	6.195	0.001	21.067	8.178
0.6 0.3 0 1 2	0.454	28.442	15.998	-0.450	14.230	7.275	-0.002	19.631	9.527
0.4 0.2 0 1 2	0.492	28.067	17.658	-0.449	13.133	8.253	-0.001	18.060	10.893
0.2 0.1 0 1 2	0.522	27.449	19.051	-0.449	12.062	9.255	0.000	16.603	12.216
0 0 0 1 2	0.542	26.735	20.315	-0.450	10.937	10.321	0.001	15.109	13.553
-0.2 -0.1 0 1 2	0.555	25.962	21.468	-0.449	9.817	11.328	0.000	13.565	14.977
-0.4 -0.2 0 1 2	0.555	24.847	22.457	-0.451	8.789	12.310	0.000	12.072	16.320
-0.6 -0.3 0 1 2	0.547	23.471	23.307	-0.450	7.645	13.390	-0.001	10.559	17.724
-0.8 -0.4 0 1 2	0.523	21.816	24.111	-0.449	6.563	14.409	-0.001	9.154	19.057
-1 -0.5 0 1 2	0.485	19.635	24.733	-0.450	5.440	15.470	0.001	7.570	20.393
-1.2 -0.6 0 1 2	0.429	16.942	25.369	-0.450	4.382	16.495	0.001	6.033	21.723
-1.4 -0.7 0 1 2	0.351	13.595	25.931	-0.450	3.230	17.528	-0.001	4.541	23.125
-1.6 -0.8 0 1 2	0.252	9.561	26.360	-0.448	2.225	18.543	-0.001	3.031	24.491
-1.8 -0.9 0 1 2	0.132	4.959	26.780	-0.449	1.097	19.576	0.000	1.490	25.792
-2 -1 0 1 2	0.001	0.014	27.205	-0.449	-0.028	20.638	-0.001	0.000	27.234

Extended Data Table 3 Coincidence degree of both evaluation and readiness processing for the task performance under the different model settings.

Models		Reaction Time				Accuracy			
		Adj. R^2	*P*	*Adj. R^2 diff.*	*$P_{adj.}$**	*Adj. R^2*	*P*	*Adj. R^2 diff.*	*$P_{adj.}$**
All Modules									
Relative-scale Models	Readiness (control W-shape)	**0.0849**	**0.0000**	**0.0119**	**0.0476**	**0.0617**	**0.0000**	0.0123	0.0506
	Evaluation (control N-shape)	**0.0252**	**0.0138**	0.0087	0.0646	0.0055	0.1607	0.0044	0.0577
Exact-scale Models	Readiness (control W-shape)	**0.0730**	**0.0000**			**0.0494**	**0.0000**		
	Evaluation (control N-shape)	**0.0165**	**0.0044**			0.0011	0.5570		
Modules Favored Relative-scale Models									
Relative-scale Models	Readiness (control W-shape)	**0.0706**	**0.0000**	**0.0124**	**0.0349**	**0.0614**	**0.0000**	**0.0159**	**0.0196**
	Evaluation (control N-shape)	**0.0164**	**0.0037**	0.0037	0.0515	0.0063	0.1245	**0.0058**	**0.0360**
Exact-scale Models	Readiness (control W-shape)	**0.0582**	**0.0000**			**0.0455**	**0.0000**		
	Evaluation (control N-shape)	**0.0127**	**0.0080**			0.0005	0.4398		
Modules Favored Exact-scale Models									
Relative-scale Models	Readiness (control W-shape)	**0.0163**	**0.0004**	**-0.0068**	**0.0463**	0.0010	0.3637	**-0.0168**	**0.0021**
	Evaluation (control N-shape)	**0.0024**	**0.0108**	**-0.0035**	**0.0073**	0.0010	0.0949	-0.0013	0.1645
Exact-scale Models	Readiness (control W-shape)	**0.0231**	**0.0000**			**0.0178**	**0.0002**		
	Evaluation (control N-shape)	**0.0059**	**0.0041**			0.0023	0.0767		

* Adjusted P values based on 10000 bootstrap; *Adj. R^2 diff.* The difference between the adjusted R square of relative-scale and exact-scale models.