Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Semantic associations and executive control predict Novelty and appropriateness of idea generation respectively and mediated by separate brain connectivity

Xueyang Wang
Southwest University
Kaixiang Zhuang
Southwest University
Jingyi Zhang
Southwest University
Robert Cortes
Danny Holzman
Li Fan
Cheng Liu
Qunlin Chen
Jiangzhou Sun
Xianrui Li
Yu Li
Qiuyang Feng
Hong Chen
Tingyong Feng
Xu Lei
Qinghua He
Adam Green
aeg58@Georgetown.edu
Georgetown University https://orcid.org/0000-0003-4178-836X
Jiang Qiu
qiuj318@swu.edu.cn
Southwest University

Article

Keywords:

Posted Date: December 27th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3691289/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Communications Biology on June 7th, 2024. See the published version at https://doi.org/10.1038/s42003-024-06405-0.

Semantic associations and executive control predict Novelty and appropriateness of idea generation respectively and mediated by separate brain connectivity

Xueyang Wang[a,b,1], Kaixiang Zhuang[a,b,1], Jingyi Zhang[a,b,1], Robert A. Cortes[d], Daniel D. Holzman[d], Li Fan[a,b], Cheng Liu[a,b], Qunlin Chen[a,b], Jiangzhou Sun[a,b], Xianrui Li[a,b], Yu Li[a,b], Qiuyang Feng[a,b], Hong Chen[a,b], Tingyong Feng[a,b], Xu Lei[a,b], Qinghua He[a,b], Adam E. Green[d, *], Jiang Qiu[a,b,c,*]

Author affiliations:
[a]Key Laboratory of Cognition and Personality (SWU), Ministry of Education, Chongqing, China;
[b]Faculty of Psychology, Southwest University, Chongqing, China;
[c]Southwest University Branch, Collaborative Innovation Center of Assessment Toward Basic Education Quality at Beijing Normal Normal University;
[d]Georgetown University, Department of Psychology.

[1]Contributed equally to this work as shared first authors.

*Corresponding author: Jiang Qiu and Adam E. Green
Email: qiuj318@swu.edu.cn; aeg58@Georgetown.edu;
Competing Interest Statement: The authors declare no competing interest.

Abstract

Novelty and appropriateness are two fundamental components of creativity. However, the way in which novelty and appropriateness are separated at behavioral and neural levels remains poorly understood. In the present study we aimed to distinguish behavioral and neural bases of novelty and appropriateness of creative idea generation. In alignment with two established theories of creative thinking, which respectively emphasize semantic association and executive control, behavioral results (n = 1,519) indicated that novelty was supported by associative ability, while appropriateness was supported by executive function. Next, employing a connectome predictive modeling (CPM) approach in resting-state fMRI data (n = 1,455), we defined two functional network-based models—dominated by interactions within default network (DMN) and by interactions within limbic network (LIM)—that respectively predicted novelty and appropriateness (i.e., cross-brain prediction). Furthermore, the generalizability and specificity of the two functional connectivity patterns were verified in additional resting-state fMRI and task-fMRI. Finally, the two functional connectivity patterns respectively mediated the relationship between semantic association/executive control and Novelty/appropriateness. These findings provide global and predictive distinctions between novelty and appropriateness in creative idea generation.

Significance Statement

As a driver of the innovation economy and a basis for AI models of idea generation, understanding how creativity works is timely. Creativity is generally considered to have two main components, novelty and appropriateness, but the empirical basis of this theoretical distinction is unclear. Here, a large behavioral dataset supported a theory-driven distinction between associative vs. executive processes. Additionally, taking a global and predictive approach, data-driven models of whole-brain functional connectivity distinguished and predicted novelty and appropriateness across brains in the same sample. Together, the data indicate that the novelty and appropriateness of creative idea generation are discernably independent, and identify distinct cognitive bases and neural network signatures.

Introduction

Creativity, defined as the ability to generate novel and appropriate ideas or products, plays a vital role in the development of various aspects of society, including scientific [1] and commercial progress [2]. A comprehensive understanding of creativity requires research that engages both novelty and appropriateness. Although researchers have made efforts to explore theoretically and empirically distinct aspects of creativity [3–10], the way in which novelty and appropriateness are distinctly instantiated at the behavioral and neural level remains poorly understood. The present research aimed to elucidate the respective behavioral and neural bases of novelty and appropriateness.

Extant evidence indicates that associative ability [11] and executive control [12–14] are important contributors to creative processes. The association theory of creative thinking emphasizes that creative thinking is depends on the ability to generate distant associations [11]. In support of the associative theory, several studies have found that individuals with higher creativity demonstrate stronger associative abilities [15–21]. The application of computational network science tools in recent studies of creativity has made it possible to more accurately quantify associative ability using semantic distance [22] (calculated by natural language processing, e.g., Word2vec). When the semantic distance between ideas or products is greater, the resulting new ideas or products tend to be more creative [23]. Semantic distance is more consistently related with novelty rather than appropriateness. Recent studies have found that novelty is positively correlated with semantic distance, with or without appropriateness as a covariate [24–26]. We thus hypothesize that associative ability plays a greater role in the novelty of creative ideas compared to the appropriateness of the ideas.

On the other hand, the executive theory of creative thinking emphasizes that creative cognition is influenced by top-down execution [12–14]. Studies have examined a range of executive functions in creativity, such as fluid intelligence [13,14,27–29], working memory [30–33], cognitive flexibility [12,34] and inhibitory control [35,36]. Inextricably tied to these “fluid” functions are the representations in memory (i.e. knowledge/experience) on which these functions operate to generate creative ideas by searching, reorganizing, and combining knowledge that is stored in semantic memory [37–42]. Previous research has shown that executive function contributes to memory retrieval [43–45]. The ability to extract appropriate information from semantic memory may help improve the appropriateness of creative ideas. Therefore, we hypothesized that the appropriateness of idea generation would be associated with executive function. Taken together, a balance of novelty and appropriateness, which may be supported by associative ability and executive function respectively, appears requisite for successful creative idea generation [46].

Functional magnetic resonance imaging (fMRI) studies have identified key regional and connectomic characteristics of creative neurocognition [4,9,47,48]. Core networks have emerged, including the default mode network (DMN), frontoparietal control network (FPCN), and salience network (SAL) [9], based on evidence that the coupling between and within these networks is reliably linked to creative ability. In addition, motor and premotor regions and cortical regions associated with semantic and

episodic memory [15,48–50] have been shown to play a role in creative cognition [48,51]. Some regional activation (e.g., in hippocampus and orbitofrontal cortex) has been associated with appropriate responding, and other regional activation (e.g., in dorsolateral prefrontal cortex and caudate) has been associated with novelty [52–54]. Yet the identified regions are frequently active across a wide range of both creative and non-creative tasks. Rather than regional markers, a whole-brain connectomic approach may afford the potential to globally and predictively distinguish the respective neural implementations of novelty and appropriateness [4,55–57]. In particular, connectome predictive modeling (CPM) [58] may allow data-driven isolation of the neural network signatures of novelty and appropriateness, and prediction of these two aspects of creativity from an individual's brain connectivity.

In the present study, we aimed to distinguish the novelty and appropriateness of creative idea generation at both the behavioral and neural levels. First, at the behavioral level, we tested a hypothesized dichotomy – based on the associative and executive theories of creativity – that associative ability supports novelty of creative idea generation and executive function supports appropriateness of creative idea generation. Next, in a resting-state fMRI dataset, we defined functional network-based models for novelty and appropriateness that achieved cross-brain prediction of novelty and appropriateness. Furthermore, the generalizability and specificity of the two functional connectivity patterns were verified in additional resting-state fMRI and task-fMRI. Finally, the two functional connectivity patterns respectively mediated the relationship between semantic association/executive control and Novelty/appropriateness.

Results

Associative ability predicts novelty while executive function predicts appropriateness

Novelty and appropriateness are two primary aspects of evaluating creativity, and they may be supported by different cognitive abilities. Here we hypothesized that novelty may be supported by associative ability while appropriateness may be supported by executive function. We tested these hypotheses using data (n = 1,519,1,023 females, age M = 21.1, SD = 0.98) from the Behavioral Brain Research Project of Chinese Personality (BBP)—associative ability was measured using the Remote Associates Test (RAT) and the Forward Flow Test (FFT) (Fig. 1A); distinct components of executive function were measured using the n-back task (measure working memory), the Stop Signal Task (measure inhibitory control) and the Number-Letter Category Switching Task (measure cognitive flexibility) (Fig. 1B). Novelty and appropriateness were measured using two AUT items (Brick and Can, Fig. 1C).

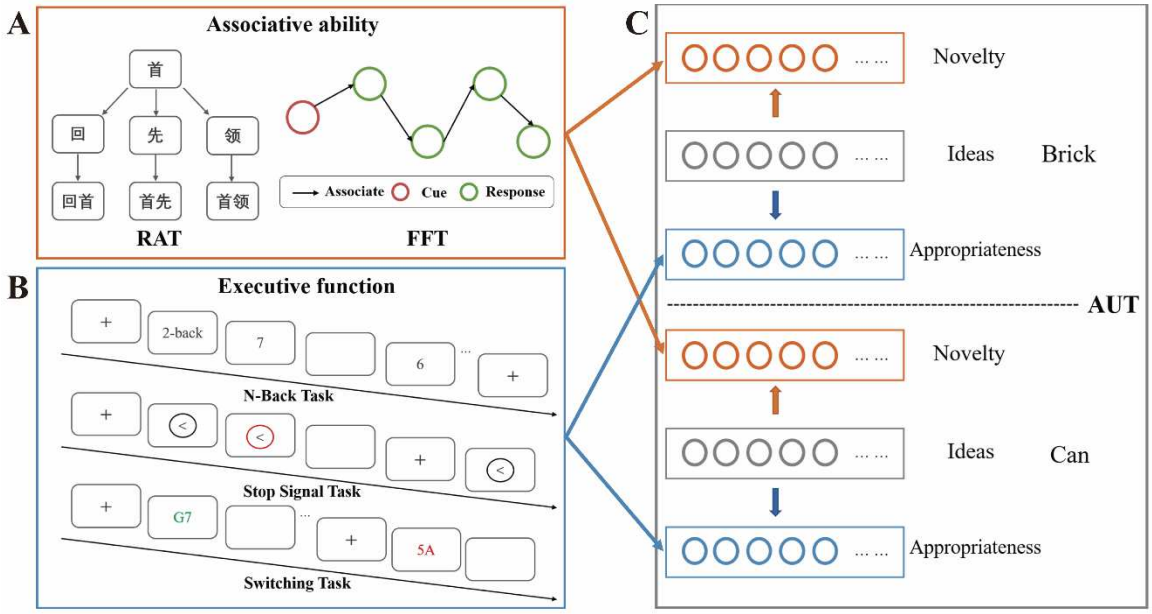


Fig. 1. Estimation of Associative ability, executive function and creativity. (**A**) One example item of RAT (left) and schematic of FFT (right). RAT: Chinese word "首" can form "回首" (Look back), "首先" (First) and "首领" (Leader) with Chinese words "回", "先" and "领" respectively. FFT: the red circle represents the cue word; the green circles represent the response words; the black arrow represents the associative process. RAT and FFT were used to measure associative ability. (**B**) Schematic of N-back Task (top), Stop Signal Task (middle) and Switching Task (also called Number-Letter Category Switching Task, bottom). These tasks were used to measure executive function. (**C**) Schematic of AUT. Brick and Can are two items of AUT. Grey circles represent the ideas generated by participants based on each item. Orange circles represent the novelty of the ideas. Blue circles represent the appropriateness of the ideas. We hypothesized that associative ability is more responsible for the novelty of the creative ideas generated and executive function is more responsible for the appropriateness of the ideas generated. RAT Remote Associates Test; FFT Forward Flow Test; AUT Alternative Uses Task.

Four trained raters evaluated the two AUT items to obtain scores of novelty and appropriateness of every participant; the higher the score, the better the novelty or appropriateness of the participant. The mean standardized reaction time (RT) of three executive function tasks was used as an indicator to represent the ability of executive function; the shorter the reaction time, the better the executive function. Meanwhile, the mean standardized accuracy (ACC) of three executive function tasks was controlled in the general linear model, due to the speed-accuracy trade-off. In addition, the standardized RAT score and the standardized semantic distance [5,18] of FFT were averaged to represent the associative ability; the higher the mean score, the better the associative ability.

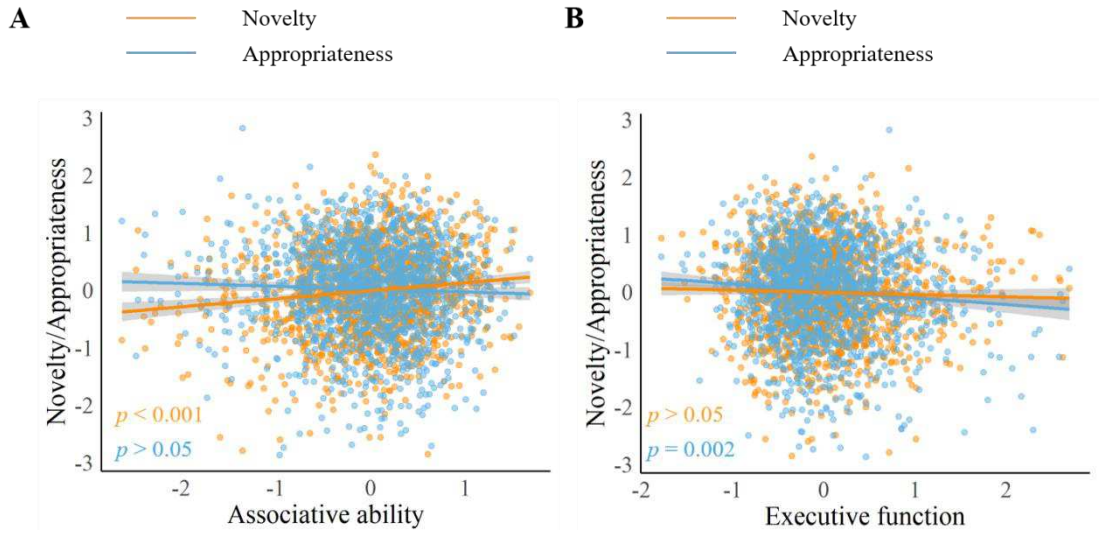


Fig 2. Results of general linear models predicting novelty and appropriateness. (**A**) The results indicated that associative ability can only predict the novelty (orange) but not appropriateness (blue). (**B**) The results indicated that executive function can only predict the appropriateness (blue) but not novelty (orange). Novelty, the standardized score of novelty; Appropriateness, the standardized score of appropriateness; Associative ability, the mean standardized RAT score and the standardized semantic distance of FFT; Executive function, the mean standardized RT of three executive function tasks.

General linear modelling (GLM) indicated that novelty was supported by associative ability, such that higher associative ability was positively associated with higher novelty ratings ($n = 1{,}519$, $\beta = 0.12$, s.e. = 0.03, $t_{1517} = 4.61$, FWE corrected $p < 0.001$, Fig. 2A orange); on the contrary, there was no significant association between novelty and executive function ($n = 1{,}519$, $\beta = -0.02$, s.e. = 0.03, $t_{1517} = -0.94$, FWE corrected $p > 0.05$, Fig. 2A blue). A second GLM indicated that appropriateness was supported by executive function, such that higher executive function was associated with higher appropriateness ratings ($n = 1{,}519$, $\beta = -0.09$, s.e. = 0.04, $t_{1517} = -3.57$, FWE corrected $p = 0.002$, Fig. 2B blue). However, there was no significant association between appropriateness and associative ability ($n = 1{,}519$, $\beta = -0.04$, s.e. = 0.03, $t_{1517} = -1.52$, FWE corrected $p > 0.05$, Fig. 2B orange). Moreover, these results remain consistent with the previous after controlling for gender, age and handedness (see Table

S1). Therefore, these results together confirmed our double-disassociation hypothesis that novelty is supported by associative ability (and not executive function) and appropriateness is supported by executive function (and not associative ability) (Fig. 1).

Two distinct connectome predictive models predict novelty and appropriateness across brains

To examine whether this double-disassociation existed at a neural level, we aimed to identify functional network-based markers of novelty and appropriateness using CPM approach [58] in resting-state fMRI (rs-fMRI) of BBP (n = 1,455, 981 females, age M = 21.1, SD = 0.98). We calculated the functional connectivity matrix based on the whole brain functional map of 300 nodes belonging to 7 functional networks [59]. We adopted relevance vector regression (RVR) to examine the predictive performance of the functional connectome on novelty and appropriateness. Within each cross-validation fold (10-fold cross-validation), we respectively identified all node pairs (edges) exhibiting suprathreshold-level ($p < 0.01$) positive and negative correlations with novelty ratings and appropriateness ratings in training set (Fig. 3A). We put these edges as input features into the RVR model for estimation to get a function that fits the behavior ratings (novelty and appropriateness) to the selected features respectively. Next, the estimated RVR model was used on the test set to obtain predicted behavioral ratings (novelty and appropriateness) (Fig. 3B). After all folds were completed, we obtained the predicted ratings for each participant. Because each random division results in different testing sets and training sets, we repeated the above prediction pipeline 150 times to generate 150 predicted ratings for each participant and further averaged these predicted ratings to obtain robust estimates. We finally correlated the predicted ratings with observed rating and then compared the correlation value with a null distribution.

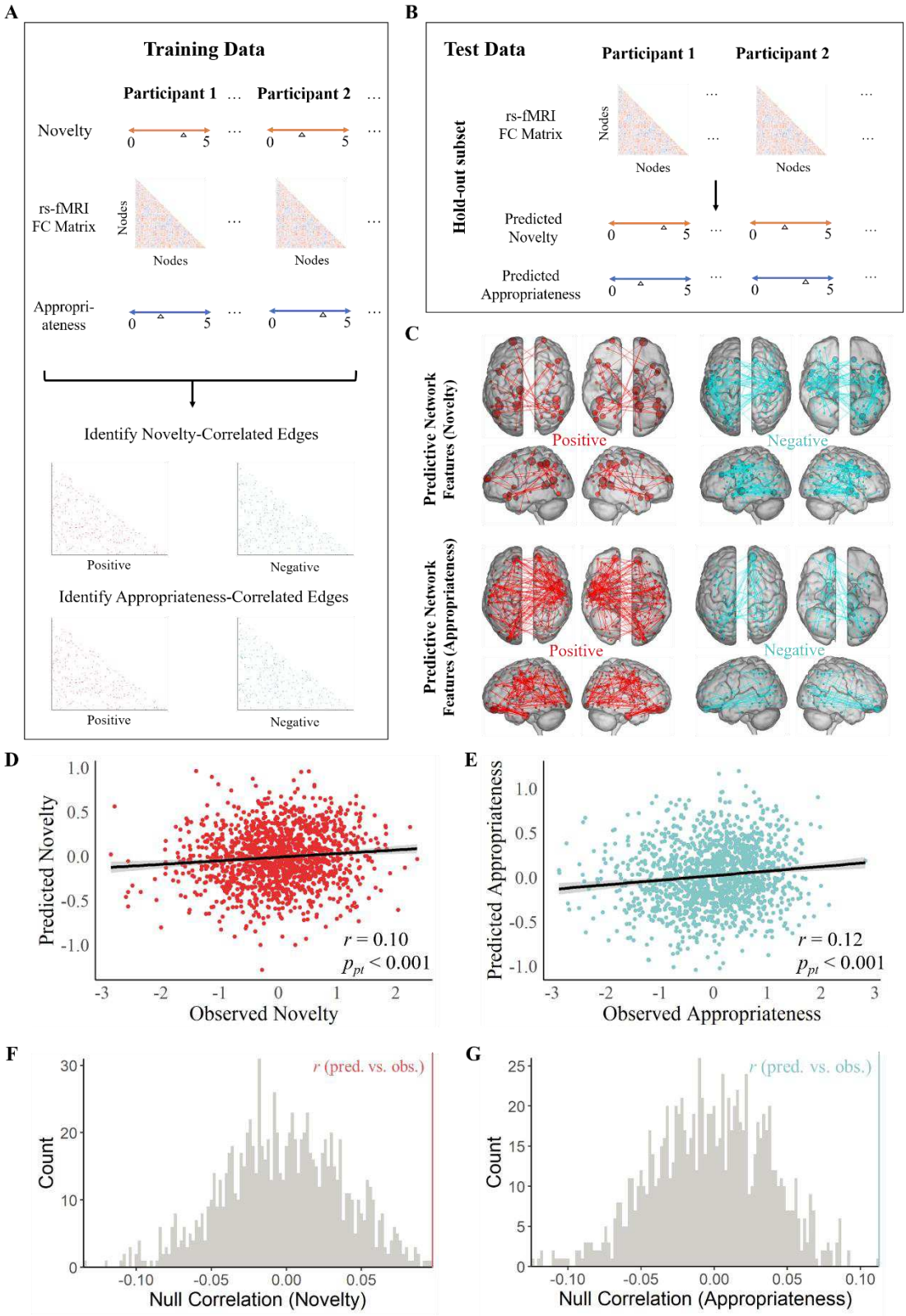


Fig. 3. Functional connectivity-based predictive modeling of trial-wise Novelty and Appropriateness fluctuations. (**A**) A schematic of the analysis pipeline. Within training data, novelty score (orange arrows), appropriateness score (blue arrows) and functional connectivity matrices (based on a 300-node whole-brain atlas) were extracted for each participant. Red and blue, respectively, indicate positive and negative functional connectivity of novelty/appropriateness.

Significant edges were identified at a threshold of $p < 0.01$ (uncorrected). (**B**) Within test data, predicted novelty ratings and predicted appropriateness ratings of the held-out subset was calculated based on the training result. (**C**) Edges strongly contributing positively(red) and negatively (blue) to the predictive model of novelty (top) and the predictive model of appropriateness (bottom). A degree threshold of 1 was applied; i.e., nodes involved in at least 1 contributing edge are displayed. (**D**) The plot shows the correlation between the predicted and observed novelty ratings based on brain connectivity for the significant predictions. (**E**) Same as (**D**), except for the predicted and observed appropriateness ratings. (**F**) Correlation value (indicated with red line) between predicted and observed novelty ratings was compared with a null distribution of r values derived from 1000 permutations of shuffled functional connectivity matrix. (**G**) Same as (**F**), except for the predicted and observed appropriateness ratings.

The permutation results showed that CPM of novelty ($r = 0.099$, $p_{pt} < 0.001$) and CPM of appropriateness ($r = 0.121$, $p_{pt} < 0.001$) were both effective (Fig. 3D – 3G). This provided evidence that our functional network model predicted novelty and appropriateness within individuals at the within-dataset level (i.e., establishing internal validation) and with an overall effect size that was on par with that typically found for functional connectivity-based prediction of self-report outcomes [60,61]. The edges contributing to the models included 296 and 430 edges, respectively, positively and negatively associated with novelty and appropriateness respectively (hereafter referred to as "novelty-CPM" and "Appropriateness-CPM" masks). These edges of Novelty-CPM and Appropriateness-CPM were both distributed widely throughout the brain, with Novelty-CPM high-degree nodes (i.e., nodes involved in multiple contributing edges) situated in temporal, prefrontal, parietal and occipital cortices and with Appropriateness-CPM high-degree nodes situated in prefrontal, temporal, and parietal cortices (Fig. 3C).

Validity and specificity of the Novelty-CPM and Appropriateness-CPM

We performed a series of confirmatory analyzes to assess the validity and specificity of the novelty CPM and appropriateness CPM. In order to avoid the arbitrariness of a single threshold, we additionally applied two common thresholds (uncorrected p-value: 0.05 and 0.005) for feature selection. We found that both Novelty-CPM ($r_{0.05} = 0.080$, $p_{pt} = 0.011$; $r_{0.005} = 0.094$, $p_{pt} = 0.005$) and Appropriateness-CPM ($r_{0.05} = 0.125$, $p_{pt} < 0.001$; $r_{0.005} = 0.109$, $p_{pt} = 0.001$) remain significant under these two thresholds. Moreover, although fMRI preprocessing steps can reduce the impact of noise on functional connectivity, frame-wise head motion can still affect the relationship between functional connectivity and behavior [62,63]. Therefore, we re-conducted the CPM model under three thresholds (uncorrected p-value: 0.05, 0.01 and 0.005) while controlling for mean framewise head motion and common demographic variables (age, gender and handedness). We also found that both Novelty-CPM ($r_{0.05} = 0.075$, $p_{pt} = 0.013$; $r_{0.01} = 0.097$, $p_{pt} = 0.004$; $r_{0.005} = 0.089$, $p_{pt} = 0.007$) and Appropriateness-CPM ($r_{0.05} = 0.127$ $p_{pt} < 0.001$; $r_{0.01} = 0.119$, $p_{pt} = < 0.001$; $r_{0.005} = 0.114$, $p_{pt} < 0.001$) remain significant when using partial correlations for feature selection, controlling for mean framewise head motion, age, gender and handedness.

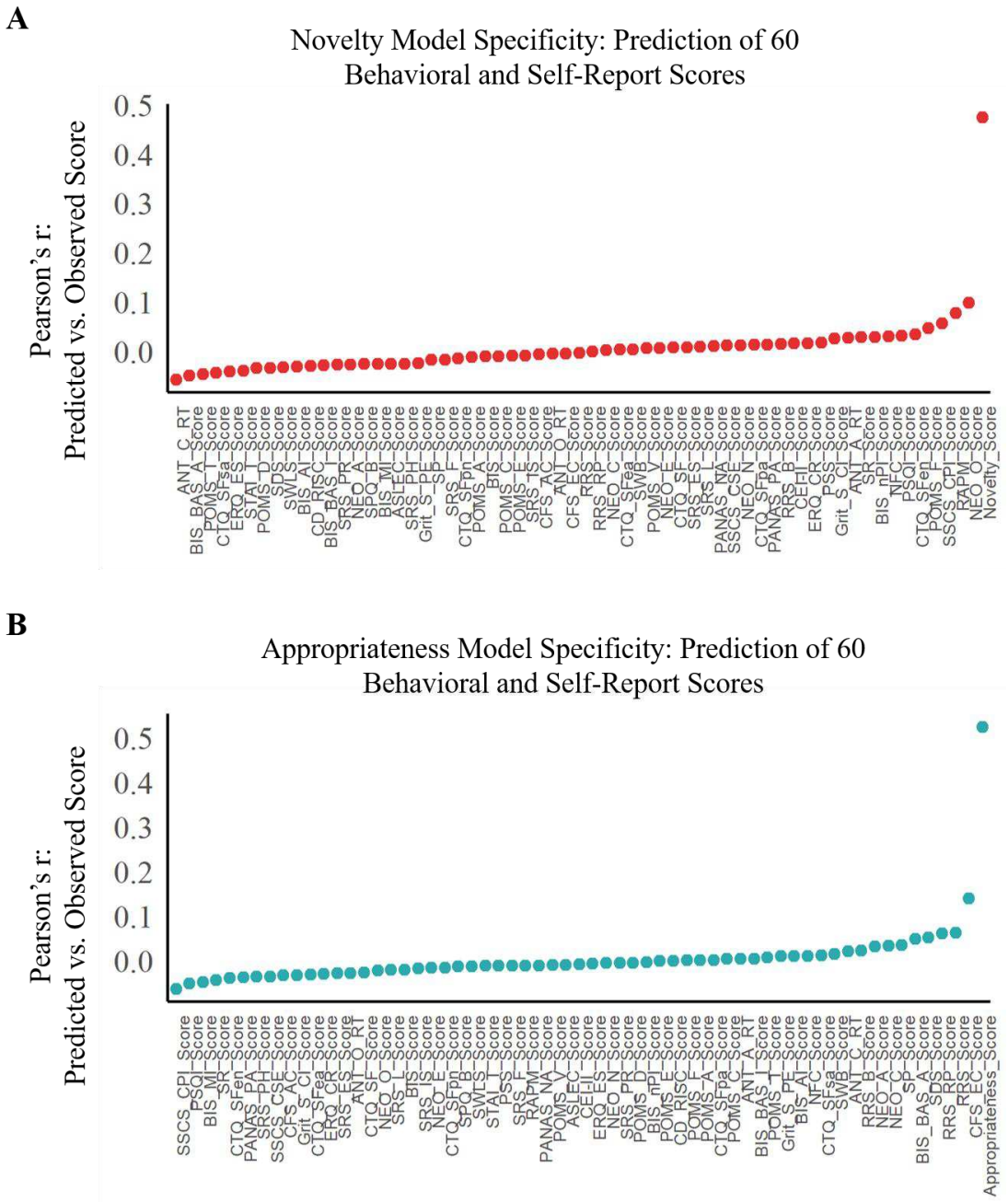


Fig. 4. The specific functional neuroanatomical basis for novelty and appropriateness derived from rs-fMRI (BBP dataset). (**A**) Correlations between Novelty-CPM prediction and 60 behavioral and self-report outcomes in the BBP dataset. Among all outcomes, novelty (denoted as "Novelty_Score") showed the highest correlation with model prediction. (**B**) Same as (**A**), except for the Appropriateness-CPM and appropriateness (denoted as "Appropriateness_Score"). See Supplementary Methods for a phenotype legend of labels shown. CPM connectome-based predictive model.

To illustrate the specificity of functional connectivity patterns that predicted observed novelty score and observed appropriateness score, we conducted the Pearson's correlation between Novelty/Appropriateness-CPM predictions from rs-

fMRI and a comprehensive battery of 60 distinct behavioral and self-report individual outcomes that were obtained in addition to novelty and appropriateness ratings (see Supplementary Methods). Among this entire set of individual outcomes, novelty prediction showed the strongest correlation with observed novelty ratings (Fig. 4A), appropriateness prediction showed the strongest correlation with observed appropriateness ratings (Fig. 4B). However, inconsistencies in sample size caused by data matching between behavioral outcomes and functional connectivity matrixes from rs-fMRI may bring comparison errors, we therefore repeated the above process using effect size (η^2), which is not affected by sample size [64]. We still found that Novelty - CPM has the highest effect size on observed novelty ratings (see Table S2) and Appropriateness-CPM has the highest effect size on observed appropriateness ratings (see Table S3).

Functional neuroanatomical basis of the novelty and appropriateness networks

In order to better explain the functional neuroanatomical basis of patterns contributing to the Novelty-CPM and Appropriateness-CPM, we examined the relationships with the functional networks previously associated with novelty and appropriateness. According to Schaefer300 atlas, with each node assigned to 1 of 7 standard Yeo-Krienen networks [65], we quantified the number of Novelty-CPM and Appropriateness-CPM mask edges belonging to each intra- or inter-network pair. The highest number of edges positively correlated with novelty (Fig. 5A) derived from Default mode network (DMN) within-network and between-network connections. The top 3 network pairs contributing to positive edges were DMN-DMN, DMN-LIM and DMN-frontoparietal control network (FPCN). The DMN-DMN connections contributed the most positive edges (19.05%). The top 3 network pairs contributing to negative edges were DMN-visual network (VIS), dorsal attention network (DAN)- salience network (SAL) and DAN-FPCN (Fig. 5B). The DMN-VIS connections contributed the most positive edges (17.80%).

The highest number of edges positively correlated with appropriateness (Fig. 5C), derived from SAL within-network and between-network connections. The top 3 network pairs contributing to positive edges were limbic network (LIM) within-network, SAL-VIS and LIM-SAL. The LIM-LIM connections contributed the most positive edges (19.72%). The highest number of edges negatively correlated with appropriateness (Fig. 5D), derived from VIS within-network and between-network connections. The top 3 network pairs contributing to negative edges were VIS-LIM, VIS-FPCN and DMN-DMN. The VIS-LIM connections contributed the most positive edges (27.27%). These results collectively suggest that novelty and appropriateness were supported by different functional brain network-based markers.

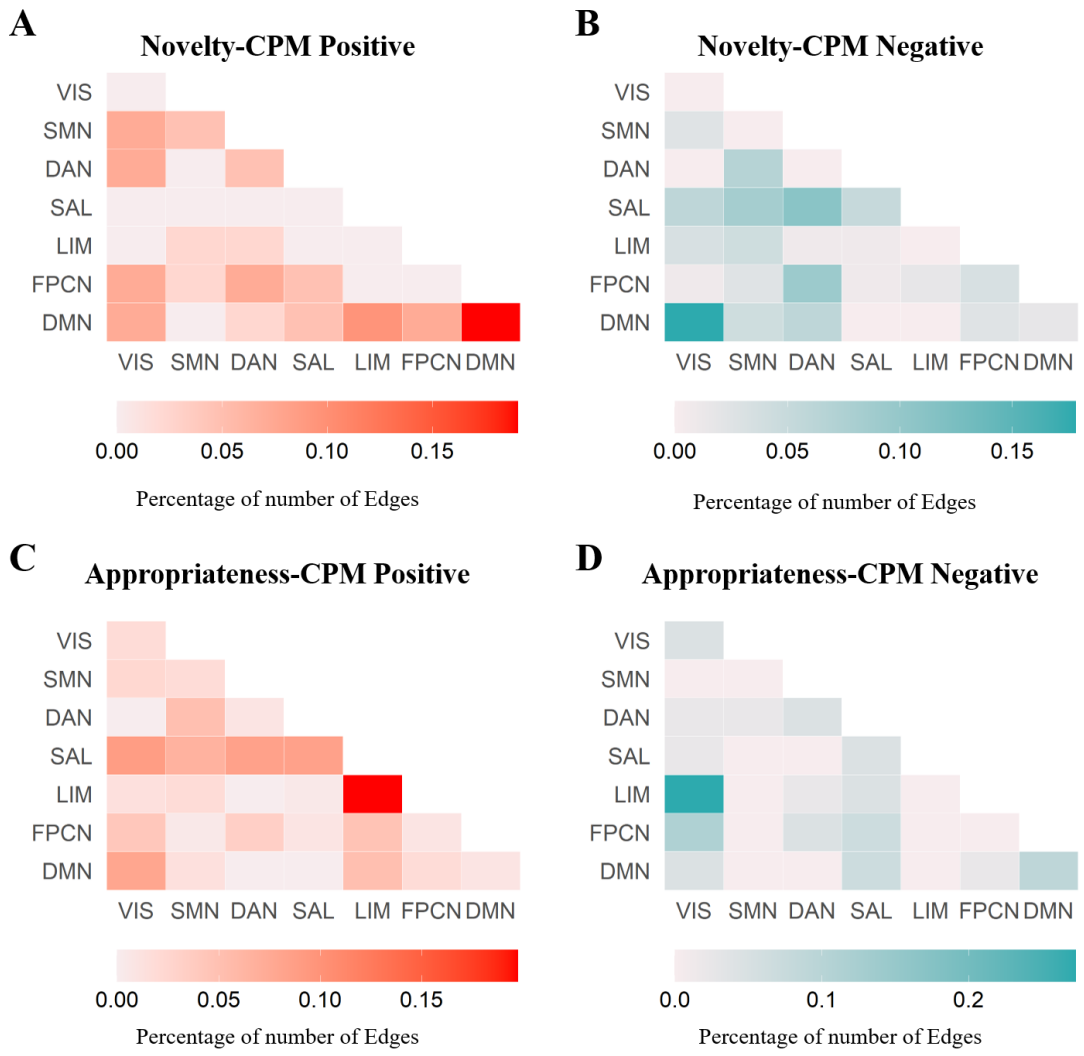


Fig. 5. Functional neuroanatomical basis of the Novelty-CPM and Appropriateness-CPM network. (**A**) The percentage of number of edges, among those within the Novelty-CPM positive mask, assigned to each within- or between-network pair based on the Schaefer300 and Yeo-Krienen 7-network atlases. (**B**) Same as (**A**), except for the Novelty-CPM negative mask. (**C**) Same as (**A**), except for the Appropriateness-CPM positive mask. (**D**) Same as (**C**), except for the Appropriateness-CPM negative mask. DAN dorsal attention network, DMN default mode network, FCPN frontoparietal control network; LIM limbic network, SAL salience network, SMN sensorimotor network, TP temporal-parietal network, VIS visual network, CPM connectome-based predictive model.

To further illustrate the robustness of Novelty-CPM and Appropriateness-CPM, we compared the results of the 6 models at the network level. As shown in Fig. 5 and Fig. S1, they are visually similar, especially in the top 1 network pairs: DMN-DMN for positive network pairs of Novelty-CPM; DMN-VIS for negative network pairs of Novelty-CPM; LIM-LIM for positive network pairs of Appropriateness-CPM; VIS-LIM for negative network pairs of Appropriateness-CPM. We next calculated the average correlation coefficients of the 6 models in the four network pairs. The results showed that positive ($r = 0.821 \pm 0.106$) and negative ($r = 0.862 \pm 0.062$) network pairs

of Novelty-CPM, positive ($r = 0.789 \pm 0.128$) and negative ($r = 0.758 \pm 0.161$) network pairs of Appropriateness-CPM between the 6 models had strong correlations. This revealed that both Novelty-CPM and Appropriateness-CPM have good robustness.

External validation #1: Applying the connectome predictive models to predict novelty and appropriateness in an independent rs-fMRI dataset

Having established the functional neuroanatomical basis of Novelty-CPM and Appropriateness-CPM based on rs-fMRI, we next explored whether the predictive models can be generalized to the prediction of novelty and appropriateness from the resting-state functional connectivity in another independent dataset (n = 46, 37 females, age M = 21.7, SD = 1.92) (hereafter referred to as "EV1 dataset"). To this end, we used the positive- and the negative- models from the resting-state functional connectivity data from BBP dataset to build the predictive model for novelty and appropriateness (using the same method as BBP dataset prediction described above). We next applied Novelty-CPM and Appropriateness-CPM to the functional connectivity of each participant from new rs-fMRI dataset to compute the predicted novelty/appropriateness ratings respectively. The model prediction was assessed by the Pearson's partial correlation between the predicted value from the model and the observed values (controlling for mean framewise head motion, age and gender). We found a significant positive correlation between novelty ratings and Novelty-CPM prediction from resting-state functional connectivity ($r = 0.26$, $p = 0.044$, one-tailed, Fig. 6A). The Appropriateness-CPM prediction of resting-state connectivity was positively significant correlated with appropriateness ratings ($r = 0.29$, $p = 0.029$, one-tailed, Fig. 6B).

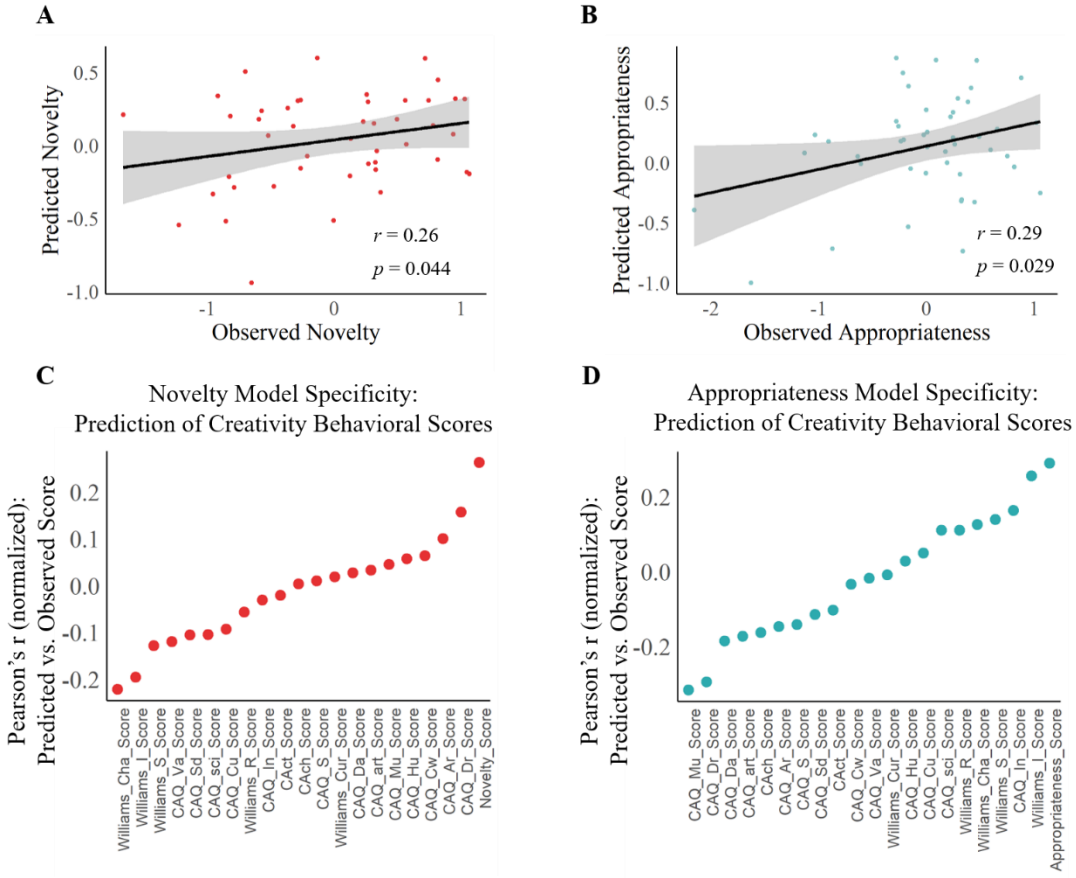


Fig. 6. The external validation and specificity of Novelty-CPM and Appropriateness-CPM on EV1 dataset. (**A**) The plot shows the correlation between the predicted and observed novelty ratings based on resting-state brain connectivity from EV1 dataset. (**B**) Same as (**A**), except for the predicted and observed appropriateness ratings. (**C**) Correlations between Novelty-CPM prediction and creativity behavioral scores in the EV1 dataset. Among all outcomes, novelty (denoted as "Novelty_Score") showed the highest correlation with model prediction. (**D**) Same as (**C**), except for the Appropriateness-CPM and appropriateness (denoted as "Appropriateness_Score"). See Supplementary Methods for a phenotype legend of labels shown. CPM connectome-based predictive model.

We have illustrated the specificity of Novelty-CPM and Appropriateness-CPM using several distinct behavioral and self-report individual outcomes in BBP dataset. However, these behaviors do not directly measure creativity. Therefore, in this dataset, we used behavioral outcomes measuring real-life creativity and creative tendencies including Creative Achievement Questionnaire (CAQ), International Creative Arts Alliance (ICAA) and Williams' Creativity Assessment Packet (WCAT) to further verify the specificity of Novelty-CPM and Appropriateness-CPM for divergent thinking. Among this entire set of individual outcomes, novelty prediction showed the strongest correlation with observed novelty score (Fig. 6C), appropriateness prediction showed the strongest correlation with observed appropriateness score (Fig. 6D). Because the data involved here had consistent sample sizes, we did not repeat this part of the analysis further using effect size.

External validation #2: Applying the connectome predictive models to predict novelty and appropriateness in an independent task-fMRI dataset

We have so far established a functional network model that carries predictive information about novelty and appropriateness, measured during rs-fMRI. In the next analysis, we focused on whether the Novelty-CPM and Appropriateness-CPM are sensitive to task performance of novelty and appropriateness respectively based on task-fMRI. We used task-fMRI of AUT from another independent dataset (n = 31, 24 females, age M = 21.8, SD = 1.88) to examine generalizability of Novelty-CPM and Appropriateness-CPM. In each trial of this paradigm (Fig. S2), an object was presented on screen (e.g., hair, see Supplementary Table 5 for details of the objects presented in each trial). In the Novelty Use (NU) condition (31s per trial), participants were asked to think of as many novel and appropriate uses as they could for each object. In the General Characteristic (GC) condition (21s per trial), participants were asked to think of as many general characteristics as they could for each object (see "Methods" for details of the stimulus presentation and procedure). Participants responded verbally using a microphone. In the analysis of task-fMRI data, only the NU condition was considered.

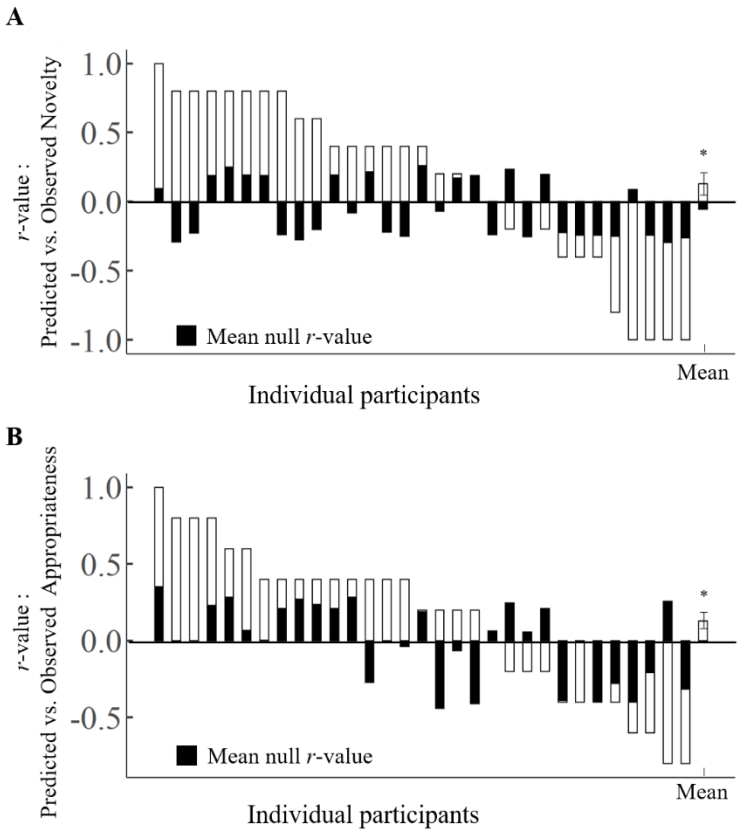


Fig. 7. The external validation of Novelty-CPM and Appropriateness-CPM on task-fMRI. (**A**) Predicted versus observed novelty ratings Spearman correlations, and mean null correlations, within each held-out participant. The last bar (far right) indicates the mean across individuals (i.e., mean across white bars; n = 31). At the group level, predicted versus observed correlations were

significantly greater than mean null correlations ($p = 0.047$, one-tailed, Wilcoxon signed-rank test). (**B**) Same as (**A**), except for the CPM prediction of appropriateness and appropriateness ratings. CPM connectome-based predictive model. *$p < 0.05$.

Refer to previous research analysis[60], we applied Novelty-CPM and Appropriateness-CPM to the functional connectivity of each block for each participant to compute the predicted novelty/appropriateness ratings for each block respectively (see "Methods" for details). We next calculated the Spearman rank correlation between predicted novelty/appropriateness and observed novelty/appropriateness respectively, and then compared the correlation value with a null distribution. Correlations between Novelty-CPM prediction (based on the rs-fMRI of BBP dataset) and observed novelty (from task-fMRI) were significantly greater than mean null values (mean $r = 0.13$, $p_{pt} = 0.047$, one-tailed, Wilcoxon signed-rank test). Moreover, Appropriateness-CPM had also been successfully verified (mean $r = 0.13$, $p_{pt} = 0.043$, one-tailed, Wilcoxon signed-rank test).

Mediation analysis

In the previous analyses, we found that novelty is supported by associative ability (and not executive function) and appropriateness is supported by executive function (and not associative ability). And we also found that Novelty-CPM and Appropriateness-CPM can predict novelty and appropriateness respectively. In a final step, we analyzed whether the relationship between novelty and associative ability is mediated by the Novelty-CPM; and whether the relationship between appropriateness and executive function is mediated by the Appropriateness-CPM. This analysis used the data from BBP dataset.

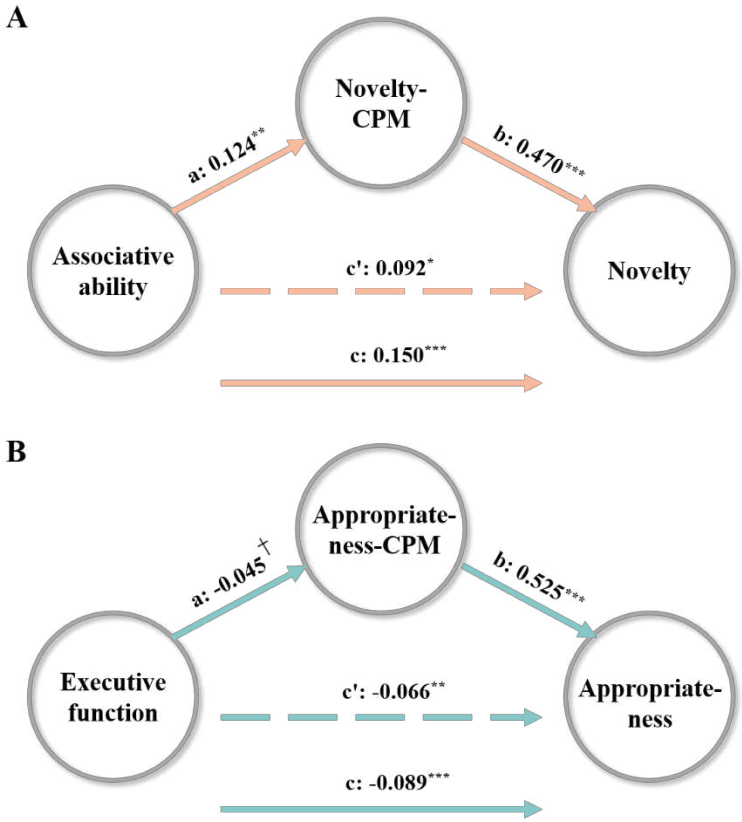


Fig. 8. Mediation analyses. Results of the mediation models are presented in path diagrams. Each diagram indicates the beta weights of the regression coefficients. The total effect is indicated by path c, the direct effect by path c′, and the indirect effect is given by the product of path a and path b. (**A**) The mediating role of Novelty-CPM on the relationship between the associative ability and novelty ratings. (**B**) The mediating role of Appropriateness-CPM on the relationship between the executive function and appropriateness ratings. †$p = 0.098$, *$p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

We explored the mediating role of Novelty-CPM on the relationship between the associative ability and novelty ratings (Fig. 8A). As shown in the previous analyses, the regression coefficient between novelty ratings and Novelty-CPM was statistically significant ($\beta = 0.470$, $p < 0.001$), as was the regression coefficient between associative ability and Novelty-CPM ($\beta = 0.124$, $p < 0.01$). The total effect and the direct effect were statistically significant ($\beta = 0.150$, $p < 0.001$; $\beta = 0.092$, $p < 0.05$). We tested the significance of the indirect effect using a bootstrapping method. The bootstrapped indirect effect was $(0.124) \times (0.470) = 0.058$, and the 95% confidence interval ranged from 0.019 to 0.100. Thus, the indirect effect was statistically significant. Hence, Novelty-CPM mediated the relationship between the associative ability and novelty ratings.

We next explored the mediating role of Appropriateness-CPM on the relationship between the executive function and appropriateness ratings (Fig. 8B). As shown in the previous analyses, the regression coefficient between appropriateness ratings and Appropriateness-CPM was statistically significant ($\beta = 0.525$, $p < 0.001$), and the

regression coefficient between executive function and Appropriateness-CPM was marginally significant (β = -0.045, p = 0.098). The total effect and the direct effect were statistically significant (β = -0.089, p < 0.001; β = 0.-066, p < 0.01). We tested the significance of the indirect effect using a bootstrapping method. The bootstrapped indirect effect was (-0.045) × (0.525) = -0.024. Although the mediation model is not significant at the 95% confidence interval, it is significant at the 90% confidence interval (ranged from -0.047 to -0.001).

Discussion

The present study provides evidence for separating the novelty and appropriateness of creative idea generation at both the behavioral and neural level. Previous work has associated localized differences in brain activity with the generation of novel and appropriate ideas by manipulating conditions of idea generation [52–54]. The present study takes a global and predictive approach, developing data-driven models of whole-brain functional connectivity that distinguished and predicted novelty and appropriateness across brains and between resting state and task-based fMRI. Additionally, a separate behavioral component of the present study supported a theory-driven framework for distinguishing divergence and convergence based on associative vs. executive processes. Together, the present data indicate that the novelty and appropriateness of creative idea generation are discernably independent, and identify distinct cognitive bases and neural network signatures for these two primary components of creativity.

Our results support the role of associate ability as a basis for novelty in creative processes (Fig. 2). This finding aligns with the associative theory of creativity, which holds that novelty of creative ideas/products represents an ability to produce distant associations [11]. Although novelty is an important aspect of creativity, solely pursuing novelty in creative activities is not practical or realistic. Creative outcomes should be balanced in terms of novelty and appropriateness. Just as the proverb, "people can't make bricks without straw" suggests, creative activities rely on the individual's existing knowledge and experience [6]. For a specific creative activity, individuals first must possess certain knowledge and experience related to the current creative activity before they can generate an appropriate product or idea. At the same time, the knowledge and experience stored in the individual's mind is so rich and complex that the individual needs a matching cognitive ability (such as executive function) to be able to extract and process the knowledge and experience related to the current creative activity. This may explain why appropriateness is supported by executive function (Fig. 2). Specifically, cognitive flexibility is responsible for screening as much information as possible; inhibitory control is responsible for excluding irrelevant information; and working memory is responsible for providing storage space for the above information processing activities.

Using CPM, we uncovered distinct whole-brain networks associated with novelty and appropriateness. Network edges identified for both Novelty-CPM and Appropriateness-CPM were widely distributed throughout the brain, with Novelty-CPM high-degree nodes (i.e., nodes involved in multiple contributing edges) situated in temporal, prefrontal, parietal and occipital cortices, and with Appropriateness-CPM high-degree nodes situated in prefrontal, temporal, and parietal cortices (Fig. 3C). Our findings show that the distributed nature of interactions within and between different brain networks was critical for the generalizability of Novelty-CPM predictions and Appropriateness-CPM predictions. Notably, the connectomic signatures of novelty and appropriateness exhibited both distinct and overlapping components. Specifically, DMN within-network (positive support) and DMN-VIS (negative support) distinctly supported novelty, LIM within-network (positive support) and LIM-VIS (negative

support) connections distinctly supported appropriateness (Fig. 5). This illustrates that the generation of novel ideas is more supported by DMN, while the generation of appropriate ideas is more supported by LIM. Whether generating novel or appropriate ideas, VIS is an overlapping component, and VIS is less synchronized with other network activities (DMN and LIM).

The distinct functional connectivity of novelty is remarkably similar to patterns of functional connectivity reported in prior researches of creative cognition [4,9]. Specifically, we found that the greatest number of significantly correlated interactions occurred within hub of one large-scale brain network: DMN (Fig. 5A). These regions revealed increased functional connectivity in a previous fMRI study of divergent thinking [66]. Notably, neuroimaging studies have also reported that artistic performance, music composition, and real-life creative achievement are also associated with functional coupling within region in this network [7,67–70], which may reflect a domain-general mechanism of creative information processing.

For the distinct neuroanatomy of appropriateness, the distributed nature of interactions within LIM plays key roles in the generation of appropriate ideas (Fig. 5C). These findings extend previous findings on the neural basis of appropriateness [52–54]. LIM contains two brain areas temporal pole (TP) and orbitofrontal cortex (OFC) [65]. Multiple functions have been ascribed to the OFC including mediating context specific responding [71], encoding contingencies in a flexible manner, encoding value, encoding inferred value, inhibiting responses, learning changes in contingency, emotional appraisal [72], altering behavior through somatic markers, driving social behavior, and representing state spaces [73]. Specific functions have been ascribed to subregions of the OFC. The lateral OFC has been proposed to reflect potential choice value, enabling fictive(counterfactual) prediction errors to potentially mediate switching choices during reversal, extinction and devaluation [74]. On the other hand, TP is a part of the temporal lobe, which is mainly responsible for processing speech information and also has the function of assigning meaning to sound information, especially the left temporal pole [75]. These references are consistent with the process of assessing appropriateness in idea generation [76]. This may indicate that during the process of generation of creative ideas, individuals evaluate the value of the ideas generated to ensure their appropriateness.

Both in Novelty-CPM and Appropriateness-CPM negative mask, our results found a dominant role for the VIS. One explanation for these results is that highly creative people are more likely to engage in internally directed cognition in the absence of external tasks [77,78]. In other words, individuals who are able to generate more novel and appropriate ideas are better able to control themselves from being disturbed by external stimuli, which is manifested in the reduction of synchrony between VIS and other brain network activities.

Taken together, our findings distinguish, for the first time, novelty and appropriateness of idea generation in both cognitive mechanisms and brain connectivity characteristics, greatly improving our understanding of the cognitive and neural basis of creativity. In future research, we should pay more attention to the impact of the interaction of novelty and appropriateness on creative ideas generation.

Materials and Methods

Participants

Behavioral and neuroimaging data were obtained at Southwest University. All participants provided written informed consent and received payment for their time and task participation, and the research protocol was approved by the ethics committee of the review committee of the Brain Imaging Center of Southwest University.

Three independent datasets were included in this study. The first dataset comprised individuals in the Behavioral Brain Research Project of Chinese Personality (BBP), from which behavioral and rs-fMRI data were obtained for the present study. In this dataset, we first analyzed the behavior data to prove that novelty is supported by associative ability while appropriateness is supported by executive function. Therefore, only participants who performed Alternate Uses Task (AUT), two associative ability tasks (Remote Associates Test (RAT) and Forward Flow Test (FFT)) and three executive function tasks (n-Back Task, Stop Signal Task and Number-Letter Category Switching Task) were eligible to participate (n = 1,535). Participants with outliers (defined as z-scores of all values of behavioral tests exceeding 3 or −3) were excluded from this analysis. This analysis included 1,519 participants finally (mean age 21.1 ± 0.98 years; 1,023 females). Next, we identified functional network-based markers of novelty and appropriateness using Connectome-based Predictive Model (CPM) approach in resting-state fMRI (rs-fMRI) of BBP. Therefore, only participants who underwent resting-state scans and performed Alternate Uses Task (AUT) were eligible to participate (n = 1,547). Participants with mean overall frame-wise displacement (FD) of >0.20 mm (based on the Power method [79]) during rs-fMRI were excluded from this analysis. This analysis included 1,455 participants finally (mean age 21.1 ± 0.98 years; 981 females).

The second dataset was used to verify the generality of the predictive models derived from BBP dataset. This dataset is called EV1 dataset in this study. This dataset included a total of 46 participants (mean age 21.7 ± 1.92 years; 37 females).

The third dataset, a task-fMRI dataset, was used to verify whether the predictive models derived from the BBP dataset can be generalized to data with different modalities. This dataset is called EV2 dataset in this study. This dataset included a total of 55 participants and 24 participants were excluded due to not meeting the analysis conditions (see Methods "Task-fMRI of Alternative Uses Task" for details). This analysis included 31 participants finally (mean age 21.8 ± 1.88 years; 24 females).

Assessment of divergent thinking

Alternative Uses Task (AUT)

In the BBP implementation of the AUT, participants were asked to generate as many novel and appropriate uses as they could think of for a brick and a can within a three minute time frame [80]. Each object (e.g., brick) was presented on the screen for three minutes. When participants came up with an idea, they were instructed to press the "1" button first and then to type that idea. This was repeated until the end of the 3-minute period. E-prime 2.0 was used to display this task.

Based on the methodologies used in previous studies [80,81], four previously trained raters evaluated the AUT responses. All responses of the two items were assessed on four aspects of creativity: fluency, which refers to the number of responses for each object; flexibility, which measures the number of distinct categories into which the responses can be classified; novelty, which measures the originality of each response, using a 5-point scale where a "1" is very unoriginal and a "5" is very original; and appropriateness, which assesses how suitable each response is for the particular object (i.e., whether the suggested use is meaningful or nonsensical) using a 5-point scale where a "1" is very unsuitable and a "5" is very suitable. In addition to the standard scoring system, raters were asked to score novelty based on their own perception of originality and unusualness [82,83]. An analysis of the rater agreement showed good inter-rater reliability, ranging from 0.728 to 0.984.

In the EV1, except that the participants did not use E-prime for testing, the rest was completely consistent with BBP. An analysis of the rater agreement showed good inter-rater reliability, ranging from 0.736 to 0.919.

In our study, only novelty and appropriateness ratings were included in the analysis. We averaged the scores of the four raters for both the novelty and appropriateness of the two objects (can and brick). We then averaged the standardized scores of novelty and appropriateness of the two items.

Product Improvement Task (PIT)

In the EV1, PIT requires an individual to create ideas to make a toy elephant more fun, enjoyable, and appealing in a 10-min time period [84]. The method of evaluation for this task is exactly the same as the AUT. An analysis of the rater agreement showed good inter-rater reliability, ranging from 0.683 to 0.791. In our study, only novelty and appropriateness ratings were included in the analysis. We averaged the ratings of the four raters for both the novelty and appropriateness of the PIT.

In the EV1 dataset, the average ratings of AUT and PIT (novelty and appropriateness) were used for external validation of Novelty-CPM and Appropriateness-CPM as an assessment of divergent thinking.

Assessment of associative ability

Remote Associates Test (RAT)

This task was used to measure associative ability. The RAT consisted of 25 trials. Each trial began with a central fixation cross for 1000ms, then three Chinese words were presented. Participants were instructed to think of a fourth word that can form three new words when combined, respectively, with each of the three presented words. Participants were given 20s per trial. When participants thought of the correct word, they were asked to press the "1" button and then to type their answer. For example, if presented with the Chinese words "回", "先" and "领", the word "首" would be the correct response, as it forms "回首" (Look back), "首先" (First) and "首领" (Leader). Trials in which participants did not press the "1" button within 20s were judged as

incorrect. RAT score was defined as the number of correct trials. This task was presented on a PC running E-prime 2.0 software.

Forward Flow Test (FFT)

The FFT was also used to measure associative ability. In the FFT, a cue word was presented first and participants were asked to write down the first thought they could think of related to the cue word. Each subsequent thought should be based on the previous thought (e.g., cake: "birthday, candle, lighting, . . ."). There were two items (cake and snowflake) for this task and participants were instructed to generate as many as many words as possible in one minute for each item.

We used semantic distance [5,18] to evaluate associative ability. To calculate the semantic distance of FFT, our approach used word2vec [15], one of a number of natural language processing technologies in which patterns of word co-occurrence are used to construct a high-dimensional semantic space (300 dimensions in this study). This method utilizes a large number of natural language corpora, which are divided into many discrete contexts. The corpora are used to generate a co-occurrence matrix to record the frequency of each word in each context. Then a data reduction technique is applied to the matrix so that each word is represented as a high-dimensional vector. Words used in similar contexts (and therefore assumed to have related meanings) are assigned similar vectors. The word similarity obtained in this way is a powerful predictor of human judgment on semantic relevance and human performance in a series of tasks [85,86]. We used the followed equation to calculate the semantic distance of FFT for every participant after transforming words into 300-dimensional numerical vectors; where D is semantic distance between thoughts and n is the total number of thoughts within a stream.

$$(\sum_{i=2}^{n} \frac{\sum_{j=1}^{i-1} D_{i,j}}{i-1})/(\mathrm{n}-1)$$

Finally, we averaged the standardized semantic distance of the two items.

The standardized RAT score and the standardized semantic distance of FFT were then averaged to represent the associative ability; the higher the mean score, the better the associative ability.

Assessment of executive function

n-Back Task

This task was used to measure working memory in executive function. We used only one condition (2-back) for the letter n-back task. Participants were asked to detect whether the current item had flashed two items earlier in the sequence. Participants were instructed to press "F" when the current item matched two items earlier and to otherwise press "J". Participants completed 90 trials. Each trial lasted 3000ms and the letter for each trial was presented for 750ms. The Psychophysics Toolbox (http://psychtoolbox.org/) for MATLAB was used to display the stimuli.

Mean RT was calculated after excluding the trials with no response and incorrect response. ACC was also calculated after excluding the trials with no response.

Stop Signal Task

The stop signal task was used to measure inhibitory control in executive function. This task consisted of 4 blocks, each of which contained 16 stop trials and 48 go trials. Stop trials and go trials were randomly presented. Each trial began with a central fixation cross for 500ms, then a black arrow pointing left or right in a black circle was displayed on the screen. The arrow lasted for 1000ms for each trial. For the go trials, participants were asked to respond as fast and accurately as possible by pressing the "F" (for left arrow) or "J" (for right arrow) button within 1000ms. For the stop trials, the black circle turned red after a period of time. This period of time is the stop-signal delay (SSD), which was dynamically adjusted according to the participant's response in this task. Participants were asked to withhold the response they already initiated. The initial SSD value is divided into four ladders: Ladder 1 = 140ms, Ladder 2 = 180ms, Ladder 3 = 220ms, Ladder 4 = 260ms. The SSD of each participant's first stop trial was Ladder 3 (220ms). If the participant successfully inhibited on a stop trial, inhibition was made more difficult on a subsequent stop trial by raising the ladder; if the participant did not successfully inhibit, inhibition was made easier by dropping the ladder. Moreover, when the SSD was at Ladder 4 and the participant could continue to successfully inhibit, SSD continued to increase on by 60ms each time; when the SSD was at Ladder 1 and the participant continued to unsuccessfully inhibit, SSD would continue to decrease on by 60ms each time. The Psychophysics Toolbox (http://psychtoolbox.org/) for MATLAB was used to display the stimuli.

The stop signal reaction time (SSRT) denotes the latency of the stop process and it is the most important dependent variable in the task. The SSRT can be measured from the observed distribution of RTs in no-stop signal trials in combination with the inhibition function [87,88]. In this study, we calculated SSRTs for each SSD using the integration method and one overall SSRT was calculated by averaging all the SSRTs acquired from SSDs in our task [87–89]. ACC was also calculated: the number of correctly judged trials in go trials plus the number of successfully stopped trials in stop trials divided by the total number of trials.

Number-Letter Category Switching Task

The switching task was used to measure cognitive flexibility in executive function. This task consisted of 40 repetition trials and 40 switching trials. Each trial began with a central fixation cross for 150ms, then a number-letter or letter-number pair was presented for 2000ms. Participants were asked to indicate whether the number was odd or even when the letter was in red or to indicate whether the letter was a consonant or a vowel when the letter was in green. Participants were asked to respond as quickly and accurately as possible by pressing the "F" (for odd and consonant) or "J" (for even and vowel) button within 2000ms after the onset of a stimulus. If the judgment type of the current trial remained the same as the previous one, a trial (after the first one) was defined as a repetition trial, otherwise it was a switching trial. The two conditions were

presented in a pseudorandom order with no more than four repeats. The Psychophysics Toolbox (http://psychtoolbox.org/) for MATLAB was used to display the stimuli. The RT of this task was defined as the mean RT of switching trials minus the mean RT of repetition trials. The ACC was defined as the average of ACC in switching trials and ACC in repetition trials.

The standardized RT of the three executive function tasks was also averaged as an indicator to represent the ability of executive function; the shorter the reaction time, the better the executive function. The standardized ACC of three executive function tasks was also averaged.

Image acquisition and preprocessing

All the functional and structural data were obtained using a 3T SIEMENS PRISMA scanner (Erlangen, Germany) at the Brain Imaging Center of Southwest University. For both task fMRI and rs-fMRI, the functional image data were obtained using a multiband T2*-sensitive gradient-recalled single-shot echo-planar imaging pulse sequence: repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle (FA) = 90°, field of view (FOV) = 224 × 224 mm^2, slices = 62, thickness = 2.0 mm, and voxel size = 2.0 × 2.0 × 2.0 mm^3. High-resolution, three-dimensional T1-weighted structural images were obtained using a magnetization prepared rapid acquisition gradient-echo (MPRAGE) sequence: TR = 2530 ms, TE = 2.98 ms, FA = 7°, slices = 192, FOV = 256 × 256 mm^2, thickness = 1.0 mm, and voxel size = 0.5 × 0.5 × 1.0 mm^3.

For BBP, preprocessing was performed in SPM12, including slice time correction, motion correction, coregistration and affine transformation of the functional volumes to a template brain (MNI).

For EV1 and EV2 datasets, FMRIPrep [90,91] based on Nipype [92] was used to preprocess the functional image data with the following parameters. Each functional image data was slice-time corrected with 3dTshift (AFNI) and then resampled to their native space by applying a single, composited transform to correct for head-motion. The BOLD time-series were resampled to the MNI152NLin2009cAsym template. Using the implementation of Nipype (version 1.15.1), the frame-wise displacement was calculated for each functional image. Then, data were filtered temporally with a nonlinear high-pass filter with a 128-s cutoff.

After preprocessing, smoothing of fMRI data was performed using spatial convolution with a Gaussian kernel of 6 mm full-width half-maximum (FWHM) for all three datasets. We next performed fMRI denoising for all three datasets based on linear regression of the following parameters from each voxel: (a) 5 noise components each from minimally-eroded WM and CSF (one-node binary erosion of nodes with values above 50% in posterior probability maps), respectively, based on aCompCor procedures [93,94]; (b) 24 motion parameters (3 translation, 3 rotation, and associated first-order derivatives and their squares); (c) all outlier frames identified within participants; and (d) linear BOLD signal trend within session. In a separate step after nuisance regression [95], data were then temporally filtered with a bandpass of 0.008–0.09 Hz.

Task-fMRI of Alternative Uses Task (EV2 dataset)

During task fMRI runs, participants were asked to think of as many novel and appropriate uses in Novelty Use (NU) condition and to think of as many general characteristics in General Characteristic (GC) condition according to an object within the fixed time. There were 11 trials under NU condition and GC condition respectively. These 22 trials were divided into 4 blocks, the first block included 2 NU trials and 2 GC trials and block 2-4 all included 3 NU trials and 3 GC trials respectively. Trials were randomly presented in each block. The objects presented in 22 trials were different (see Table S4).

Unlike previous AUT implementations, participants were required to respond verbally during the fMRI scanning. In each trial, the condition (NU/GC) was presented on the screen first; and then cue word was continuously presented to the participants, during which time participants spoke as many novel and unusual uses of the presented cue word in NU condition and as many general characteristics of the presented cue word in GC condition. The cue word presented 31s in NU condition while 21s in GC condition.

The E-prime 2.0 was used to display the stimuli and to synchronize stimulus onset with MRI data acquisition. Audio was delivered via in-ear headphones. Participants' speech was also recorded using E-prime 2.0. The Adobe Audition was used to process the recording to obtain the responses of the participants.

Five trained raters were invited to evaluated the responses of the participants of this task. The NU responses were assessed through four aspects of creativity: fluency, flexibility, novelty and appropriateness (completely consistent with the previous "Alternative Uses Task"). The GC responses were assessed through three aspects: fluency, flexibility and appropriateness. The raters were asked to score responses beyond the scoring system according to their own perception of novelty and unusualness [82,83]. An analysis of the rater agreement showed good inter-rater reliability, with Cronbach alpha scores ranging from 0.861 to 0.950.

In this study, only the novelty and appropriateness scores of NU and image data of NU were used and the novelty and appropriateness scores were standardized (z-transformed). Since not all participants responded in the 11 NU trials, we only reserved the participants who responded in all NU trials to ensure consistency in the number of trials (24 participants were excluded for this reason).

Functional connectivity feature extraction

Within each dataset, we extracted the preprocessed BOLD time series from the mean across all voxels within each node defined based on previously described intrinsic functional network atlases in MNI space (Schaefer [59] atlases of 300 cortical regions). For the EV2 dataset (task-fMRI) only, these time series were extracted for each trial based on the 31-s window prior to ideas generation onset. For both rs-fMRI datasets (BBP and EV1), we also extracted time series across the whole duration (8 mins). We computed a matrix of functional connectivity values between all region pairs based on the Fisher z-transformed Pearson's correlation coefficient of time series.

Predictive modeling analysis of novelty and appropriateness

To examine whether this double-disassociation existed at a neural level, we aimed to identify functional network-based markers of novelty and appropriateness using CPM approach [58] in resting-state fMRI (rs-fMRI) of BBP. We adopted relevance vector regression (RVR) to examine the predictive performance of the functional connectome on novelty and appropriateness. RVR is a sparse kernel multiple regression method that uses Bayesian inference to obtain parsimonious solutions that generalize well and provide inference at low computational cost [96]. Notably, RVR has no algorithm-specific parameters, and no additional computational resources are required to estimate the optimal algorithm-specific parameters. By using 10-fold cross-validation (10F-CV) framework, participants were randomly divided into 10 subsets; nine folds (90% of participants) were used as the training set, and the remaining fold (10% of participants) was used as the testing set. Within each cross-validation fold (10-fold cross-validation), we respectively identified all node pairs (edges) exhibiting suprathreshold-level ($p < 0.01$) positive and negative correlations with novelty ratings and appropriateness ratings in training set. We put these edges as input features into the RVR model for estimation to get a function that fits the behavior ratings (novelty and appropriateness) to the selected features respectively. Next, the estimated RVR model was used on the test set to obtain predicted behavioral ratings (novelty and appropriateness). After all folds were completed, we obtained the predicted ratings for each participant. Because each random division results in different testing sets and training sets, we repeated the above prediction pipeline 150 times to generate 150 predicted ratings for each participant and further averaged these predicted ratings to obtain robust estimates. We finally correlated the predicted ratings with observed rating and then compared the correlation value with a null distribution.

To determine whether predicted versus observed correlations were statistically significant, we generated a distribution of null values. To do so, we repeated all of the same CPM procedures, as described, except the assignments of functional connectivity matrix were randomly permuted (1000 iterations) to obtain null correlation values to assess the significance (permutation test, $p_{pt} \leq 0.05$). Finally, we retained edges that positively or negatively correlated with novelty and appropriateness ratings in more than 80% of the cross-validation folds to form novelty-CPM and appropriateness-CPM masks, respectively.

To avoid the arbitrariness of a single threshold, we additionally applied two common thresholds (uncorrected p values: 0.05 and 0.005) for feature selection. In addition, we also used Pearson's partial correlation for feature selection under these three thresholds (uncorrected p values: 0.05, 0.01 and 0.005) while controlling for mean framewise head motion and common demographic variables (age, gender and handedness). In summary, we additionally built 5 models to verify the stability of the main results model.

The specificity of the Novelty-CPM and Appropriateness-CPM

To assess specificity of model prediction in the BBP dataset, we computed the Pearson's correlations between novelty/appropriateness-CPM predictions from rs-

fMRI of BBP and a comprehensive battery of 60 distinct behavioral and self-report individual outcomes that were obtained in addition to novelty and appropriateness scores (see Supplementary Methods for full description). Depending on data availability, these correlations were performed with sample sizes ranging from 817 to 1447 participants (see Table S2 and Table S3). However, inconsistencies in sample size caused by data matching between behavioral outcomes and functional connectivity matrixes from rs-fMRI may bring comparison errors, we therefore repeated the above process using effect size (η^2), which is not affected by sample size [64].

We also performed the same analysis on the EV1 dataset, except that the behavioral and self-report individual outcomes (see Supplementary Methods for full description) used was different. In addition, since the sample sizes of all data involved in the EV1 dataset are consistent, we did not additionally calculate effect sizes (η^2).

Analysis of functional neuroanatomical patterns contributing to the novelty-CPM and appropriateness-CPM

We visualized the edges comprising novelty-CPM and appropriateness-CPM masks (Fig. 3C) using the BioImage Suite Connectivity Visualization Tool (https://bioimagesuiteweb.github.io/webapp). To gain insight into the neuroanatomical patterns that contributed to the novelty-CPM and appropriateness-CPM, we assigned each node to 1 of 7 canonical Yeo-Krienen [65] intrinsic functional networks (Fig. 4). For these analyses, we used the Schaefer atlas of 300 cortical regions, which includes a Yeo-Krienen network label for each node [59]. To describe our results, we use the labels provided with the publicly available Schaefer atlas (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal).

Based on the Schaefer atlas, the novelty-CPM and appropriateness-CPM included 296 and 430 edges, respectively, in the positive and negative masks. We assigned each of these edges to 1 of 28 within- or between-network Yeo-Krienan pairs. To further verify the stability of the main CPM results, We also visualized the additional 5 models at the network level (Fig. S1). And we next calculated the average correlation coefficients of the 6 models in the four network pairs (positive and negative network pairs of Novelty-CPM; positive and negative network pairs of Appropriateness-CPM). Specifically, for each network pair, we separately calculated the matrix correlation between each of the 6 models (a total of 15 correlation coefficients). Finally, we average these correlation coefficients for each network pair.

Prediction of novelty and appropriateness from rs-fMRI (EV1 dataset)

In the EV1 dataset, we tested whether Novelty-CPM and Appropriateness-CPM predictions generalize to predictions of novelty and appropriateness from rs-fMRI data. To this end, we used the positive- and the negative- models from the resting-state functional connectivity data from BBP dataset to build the predictive model for novelty and appropriateness (using the same method as "Predictive modeling analysis of novelty and appropriateness" described). We then conducted Pearson's partial correlation between predicted novelty/appropriateness ratings and observed

novelty/appropriateness ratings respectively under controlling for gender, age and mean frame-wise head motion.

Prediction of novelty and appropriateness from task-fMRI (EV2 dataset)

In the EV2 dataset, we tested whether Novelty-CPM and Appropriateness-CPM predictions generalize to predictions of novelty and appropriateness from task-fMRI data. We first average the functional connectivity matrix of each trial within each block under the NU condition (4 blocks and 11 trials in total, see "Task-fMRI of Alternative Uses Task" for details). Refer to previous research analysis[60], for each participant in the EV2 dataset, we used the positive- and the negative- models from the resting-state functional connectivity data from BBP dataset to build the predictive model for novelty and appropriateness on each block of NU condition of AUT task-fMRI (using the same method as "Predictive modeling analysis of novelty and appropriateness" described). We next calculated the Spearman rank correlation between predicted novelty/appropriateness and observed novelty/appropriateness respectively for each participant. To generate null comparison values, we repeated all of the same CPM procedures, as described, except the assignments of functional connectivity matrix were randomly permuted (1000 iterations) to obtain null correlation values. At the group level, we performed a Wilcoxon signed-rank test to compare the within-participant prediction (predicted versus observed novelty and appropriateness ratings) versus the mean of the within-participant null correlation values.

Mediation analysis

To test whether the Novelty-CPM is also relevant for associative ability and whether the Appropriateness-CPM is also relevant for executive function, we ran mediation analyses. The mediation analysis [97] consisted in calculating the product of (i) the regression coefficient of the regression analysis on the independent variable (i.e., associative ability and executive function) to predict the mediator (i.e., Novelty-CPM and Appropriateness-CPM) and (ii) the regression coefficient of the regression analysis on the mediator to predict the dependent variable (i.e., novelty and appropriateness) when controlling for the independent variable. We also calculated the regression coefficient of the regression analysis on the independent variable to predict the dependent variable without controlling for the mediator (total effect) and when controlling for it (direct effect; Fig. 8). The indirect effect was calculated as the product of path a and path b. All the variables entered in the mediation analyses were normalized. We tested the significance of the indirect effect using bootstrapping method, computing unstandardized indirect effects for each 1000 bootstrapped samples, and the 90% confidence interval was computed by determining the indirect effects at the 5th and 95th percentiles. The mediation analyses were performed using the JASP 0.17.2.1.

References

1. Wuchty, S., Jones, B. F. & Uzzi, B. The increasing dominance of teams in production of knowledge. *Science* **316**, (2007).
2. Salter, A. & Gann, D. Sources of ideas for innovation in engineering design. *Research Policy* **32**, (2003).
3. Bellana, B., Mahabal, A. & Honey, C. J. Narrative thinking lingers in spontaneous thought. *Nature Communications* **13**, (2022).
4. Beaty, R. E. *et al.* Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 1087–1092 (2018).
5. Brucks, M. S. & Levav, J. Virtual communication curbs creative idea generation. *Nature* **605**, 108–112 (2022).
6. Lewis, P. A., Knoblich, G. & Poe, G. How Memory Replay in Sleep Boosts Creative Problem-Solving. *Trends in Cognitive Sciences* **22**, 491–503 (2018).
7. Ovando-Tellez, M. *et al.* Brain connectivity-based prediction of real-life creativity is mediated by semantic memory structure. *Science Advances* **8**, (2022).
8. Shofty, B. *et al.* The default network is causally linked to creative thinking. *Molecular Psychiatry* **27**, (2022).
9. Beaty, R. E., Benedek, M., Silvia, P. J. & Schacter, D. L. Creative Cognition and Brain Network Dynamics. *Trends in Cognitive Sciences* **20**, 87–95 (2016).
10. Yang, W. *et al.* Creative problem solving in knowledge-rich contexts. *Trends in Cognitive Sciences* **26**, 849–859 (2022).
11. Mednick, S. The associative basis of the creative process. *Psychological Review* **69**, 220–

232 (1962).

12. Nusbaum, E. C. & Silvia, P. J. Are intelligence and creativity really so different?. Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence* **39**, 36–45 (2011).
13. Beaty, R. E. & Silvia, P. J. Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts* **6**, (2012).
14. Silvia, P. J. & Beaty, R. E. Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence* **40**, (2012).
15. Liu, C. *et al.* Semantic association ability mediates the relationship between brain structure and human creativity. *Neuropsychologia* **151**, (2021).
16. Beaty, R. E., Zeitlen, D. C., Baker, B. S. & Kenett, Y. N. Forward flow and creative thought: Assessing associative cognition and its role in divergent thinking. *Thinking Skills and Creativity* **41**, (2021).
17. He, L. *et al.* The relation between semantic memory structure, associative abilities, and verbal and figural creativity. *Thinking and Reasoning* **27**, (2020).
18. Gray, K. *et al.* 'Forward flow': A new measure to quantify free thought and predict creativity. *American Psychologist* **74**, 539–554 (2019).
19. Rossmann, E. & Fink, A. Do creative people use shorter associative pathways? *Personality and Individual Differences* **49**, (2010).
20. Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E. & Benedek, M. The roles of associative and executive processes in creative cognition. *Memory and Cognition* **42**, 1186–1197

(2014).

21. Zhang, J. *et al.* Retrieval flexibility links to creativity: evidence from computational linguistic measure. *Cerebral Cortex* (2022) doi:10.1093/cercor/bhac392.
22. Kenett, Y. N. & Faust, M. A Semantic Network Cartography of the Creative Mind. *Trends in Cognitive Sciences* **23**, (2019).
23. Kenett, Y. N. What can quantitative measures of semantic distance tell us about creativity? *Current Opinion in Behavioral Sciences* **27**, (2019).
24. Beaty, R. E. & Johnson, D. R. Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods* **53**, (2021).
25. Heinen, D. J. P. & Johnson, D. R. Semantic distance: An automated measure of creativity that is novel and appropriate. *Psychology of Aesthetics, Creativity, and the Arts* **12**, 144–156 (2018).
26. Hass, R. W. Tracking the dynamics of divergent thinking via semantic distance: Analytic methods and theoretical implications. *Memory and Cognition* **45**, 233–244 (2017).
27. Benedek, M., Franz, F., Heene, M. & Neubauer, A. C. Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences* **53**, 480–485 (2012).
28. Jauk, E., Benedek, M. & Neubauer, A. C. The road to creative achievement: A latent variable model of ability and personality predictors. *European Journal of Personality* **28**, 95–105 (2014).
29. Frith, E. *et al.* Intelligence and creativity share a common cognitive and neural basis. *Journal of Experimental Psychology: General* **150**, 609–632 (2020).

30. de Dreu, C. K. W., Nijstad, B. A., Baas, M., Wolsink, I. & Roskes, M. Working memory benefits creative insight, musical improvisation, and original ideation through maintained task-focused attention. *Personality and Social Psychology Bulletin* **38**, 656–669 (2012).
31. Süß, H. M., Oberauer, K., Wittmann, W. W., Wilhelm, O. & Schulze, R. Working-memory capacity explains reasoning ability - And a little bit more. *Intelligence* **30**, (2002).
32. Lee, C. S. & Therriault, D. J. The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence* **41**, (2013).
33. Takeuchi, H. *et al.* Originality of divergent thinking is associated with working memory–related brain activity: Evidence from a large sample study: Originality and brain activity. *NeuroImage* **216**, (2020).
34. Ovando-Tellez, M. *et al.* An investigation of the cognitive and neural correlates of semantic memory search related to creative ability. *Communications Biology* **5**, (2022).
35. Cassotti, M., Agogué, M., Camarda, A., Houdé, O. & Borst, G. Inhibitory Control as a Core Process of Creative Problem Solving and Idea Generation from Childhood to Adulthood. *New directions for child and adolescent development* **2016**, (2016).
36. Camarda, A. *et al.* Do we need inhibitory control to be creative? Evidence from a dual-task paradigm. *Psychology of Aesthetics, Creativity, and the Arts* **12**, (2018).
37. Abraham, A. & Bubic, A. Semantic memory as the root of imagination. *Frontiers in Psychology* **6**, (2015).
38. Benedek, M., Könen, T. & Neubauer, A. C. Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts* **6**, 273–281 (2012).

39. Benedek, M. & Fink, A. Toward a neurocognitive framework of creative cognition: the role of memory, attention, and cognitive control. *Current Opinion in Behavioral Sciences* **27**, 116–122 (2019).

40. Gençer, E. & Yıldırım, Z. Top-Down Processing. in *Encyclopedia of Animal Cognition and Behavior* (eds. Vonk, J. & Shackelford, T.) 1–8 (Springer International Publishing, 2020). doi:10.1007/978-3-319-47829-6_1832-1.

41. Kenett, Y. & Thompson-Schill, S. L. Novel conceptual combination can dynamically reconfigure semantic memory networks. Preprint at https://doi.org/10.31234/osf.io/crp47 (2020).

42. Kounios, J. & Beeman, M. The cognitive neuroscience of insight. *Annual Review of Psychology* **65**, (2014).

43. Craik, F. I. M., Eftekhari, E., Bialystok, E. & Anderson, N. D. Individual differences in executive functions and retrieval efficacy in older adults. *Psychol Aging* **33**, 1105–1114 (2018).

44. Rajan, V., Cuevas, K. & Bell, M. A. The Contribution of Executive Function to Source Memory Development in Early Childhood. *Journal of Cognition and Development* **15**, (2014).

45. Guler, O. E. & Mackovichova, S. The role of executive function in autobiographical memory retrieval: does the type of cue word matter? *Memory* **27**, (2019).

46. Rastelli, C., Greco, A., De Pisapia, N. & Finocchiaro, C. Balancing novelty and appropriateness leads to creative associations in children. *PNAS Nexus* **1**, pgac273 (2022).

47. The Cambridge Handbook of the Neuroscience of Creativity. in *The Cambridge*

Handbook of the Neuroscience of Creativity (eds. Vartanian, O. & Jung, R. E.) iii–iii (Cambridge University Press, 2018).

48. Gonen-Yaacovi, G. *et al.* Rostral and caudal prefrontal contribution to creativity: A meta-analysis of functional imaging data. *Frontiers in Human Neuroscience* (2013) doi:10.3389/fnhum.2013.00465.
49. Benedek, M. *et al.* To create or to recall original ideas: Brain processes associated with the imagination of novel object uses. *Cortex* **99**, (2018).
50. Cogdell-Brooke, L. S., Sowden, P. T., Violante, I. R. & Thompson, H. E. A meta-analysis of functional magnetic resonance imaging studies of divergent thinking using activation likelihood estimation. *Human Brain Mapping* **41**, (2020).
51. Matheson, H. & Kenett, Y. The role of the motor system in generating creative thoughts. *NeuroImage* **213**, 116697 (2020).
52. Ren, J. *et al.* The function of the hippocampus and middle temporal gyrus in forming new associations and concepts during the processing of novelty and usefulness features in creative designs. *NeuroImage* **214**, (2020).
53. Huang, F., Tang, S., Sun, P. & Luo, J. Neural correlates of novelty and appropriateness processing in externally induced constraint relaxation. *NeuroImage* **172**, (2018).
54. Huang, F., Fan, J. & Luo, J. The neural basis of novelty and appropriateness in processing of creative chunk decomposition. *NeuroImage* **113**, (2015).
55. Gabrieli, J. D. E., Ghosh, S. S. & Whitfield-Gabrieli, S. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron* **85**, (2015).
56. Woo, C. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: Brain

models in translational neuroimaging. *Nature Neuroscience* **20**, (2017).

57. Yamashita, M. *et al.* A prediction model of working memory across health and psychiatric disease using whole-brain functional connectivity. *eLife* **7**, (2018).
58. Shen, X. *et al.* Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols* **12**, (2017).
59. Schaefer, A. *et al.* Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex* **28**, 3095–3114 (2018).
60. Kucyi, A. *et al.* Prediction of stimulus-independent and task-unrelated thought from functional brain networks. *Nature Communications* **12**, 1793 (2021).
61. Chen, J. *et al.* Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nat Commun* **13**, 2217 (2022).
62. Siegel, J. S. *et al.* Data Quality Influences Observed Links Between Functional Connectivity and Behavior. *Cereb Cortex* **27**, 4492–4502 (2017).
63. Power, J. D., Schlaggar, B. L. & Petersen, S. E. Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage* **105**, 536–551 (2015).
64. Kelley, K. & Preacher, K. J. On effect size. *Psychol Methods* **17**, 137–152 (2012).
65. Thomas Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106**, 1125–1165 (2011).
66. Beaty, R. E., Benedek, M., Barry Kaufman, S. & Silvia, P. J. Default and Executive Network Coupling Supports Creative Idea Production. *Scientific Reports* **5**, 10964 (2015).
67. Ellamil, M., Dobson, C., Beeman, M. & Christoff, K. Evaluative and generative modes of thought during the creative process. *NeuroImage* **59**, 1783–1794 (2012).

68. Liu, S. *et al.* Brain activity and connectivity during poetry composition: Toward a multidimensional model of the creative process. *Hum Brain Mapp* **36**, 3351–3372 (2015).

69. Pinho, A. L., de Manzano, Ö., Fransson, P., Eriksson, H. & Ullén, F. Connecting to Create: Expertise in Musical Improvisation Is Associated with Increased Functional Connectivity between Premotor and Prefrontal Areas. *J Neurosci* **34**, 6156–6163 (2014).

70. He, R. *et al.* The impact of knowledge on poetry composition: An fMRI investigation. *Brain and Language* **235**, 105202 (2022).

71. Wikenheiser, A. M. & Schoenbaum, G. Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat Rev Neurosci* **17**, 513–523 (2016).

72. Fettes, P., Schulze, L. & Downar, J. Cortico-Striatal-Thalamic Loop Circuits of the Orbitofrontal Cortex: Promising Therapeutic Targets in Psychiatric Illness. *Frontiers in Systems Neuroscience* **11**, (2017).

73. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron* **81**, 267–279 (2014).

74. Mj, T. *et al.* Neural systems for choice and valuation with counterfactual learning signals. *NeuroImage* **89**, (2014).

75. Tsapkini, K., Frangakis, C. E. & Hillis, A. E. The function of the left anterior temporal pole: evidence from acute stroke and infarct volume. *Brain* **134**, 3094 (2011).

76. Benedek, M., Beaty, R. E., Schacter, D. L. & Kenett, Y. N. The role of memory in creative ideation. *Nat Rev Psychol* **2**, 246–257 (2023).

77. W, O. *et al.* Creativity at rest: Exploring functional network connectivity of creative experts. *Network neuroscience (Cambridge, Mass.)* **7**, (2023).

78. Orwig, W., Diez, I., Vannini, P., Beaty, R. & Sepulcre, J. Creative Connections: Computational Semantic Distance Captures Individual Creativity and Resting-State Functional Connectivity. *Journal of cognitive neuroscience* **33**, 499 (2021).

79. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142–2154 (2012).

80. Wang, X., Zhuang, K., Li, Z. & Qiu, J. The functional connectivity basis of creative achievement linked with openness to experience and divergent thinking. *Biological Psychology* **168**, (2022).

81. Chen, Q. *et al.* Brain hemispheric involvement in visuospatial and verbal divergent thinking. *NeuroImage* **202**, (2019).

82. Long, H. More than appropriateness and novelty: Judges' criteria of assessing creative products in science tasks. *Thinking Skills and Creativity* **13**, 183–194 (2014).

83. Takeuchi, H. *et al.* White matter structures associated with creativity: Evidence from diffusion tensor imaging. *NeuroImage* **51**, 11–18 (2010).

84. Chen, Q. *et al.* Longitudinal alterations of frontoparietal and frontotemporal networks predict future creative cognitive ability. *Cerebral Cortex* **28**, 103–115 (2018).

85. Bullinaria, J. A. & Levy, J. P. Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods* **39**, 510–526 (2007).

86. Recchia, G. & Jones, M. N. More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods* **41**, 647–656 (2009).

87. Logan, G. D. & Cowan, W. B. On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review* **91**, 295–327 (1984).

88. Verbruggen, F. *et al.* A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife* **8**, (2019).

89. Band, G. P. H., van der Molen, M. W. & Logan, G. D. Horse-race model simulations of the stop-signal procedure. *Acta Psychologica* **112**, 105–142 (2003).

90. Gorgolewski, K. *et al.* Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in Python. *Frontiers in Neuroinformatics* **5**, (2011).

91. Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods* **16**, 111–116 (2019).

92. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341 (2014).

93. Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007).

94. Chai, X. J., Castañán, A. N., Öngür, D. & Whitfield-Gabrieli, S. Anticorrelations in resting state networks without global signal regression. *NeuroImage* **59**, 1420–1428 (2012).

95. Hallquist, M. N., Hwang, K. & Luna, B. The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *NeuroImage* **82**, (2013).

96. Tipping, M. E. Sparse bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **1**, 211–244 (2001).

97. Introduction to Mediation, Moderation, and Conditional Process Analysis: Third Edition:

A Regression-Based Approach. *Guilford Press* https://www.guilford.com/books/Introduction-to-Mediation-Moderation-and-Conditional-Process-Analysis/Andrew-Hayes/9781462549030.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.pdf