Supplemental Online Material

Supplemental Methods

Subjects. The present study, which is part of an ongoing project to investigate associations between brain imaging, cognitive function, and aging, included relevant cognitive measures and imaging data from 1221 healthy, right-handed individuals (700 men and 521 women). The mean age (± standard deviation, SD) of subjects was 20.7 ± 1.8 years (range, 18–27 years). All subjects were university students, postgraduates, or university graduates of less than one year's standing. All subjects had normal vision and none had a history of neurological or psychiatric illness. Handedness was evaluated using the Edinburgh Handedness Inventory (Oldfield, 1971). Written informed consent was obtained from each subject prior to their participation. For non-adult subjects, written informed consent was obtained from their parents or guardians. This study was approved by the Ethics Committee of Tohoku University.

Subjects were instructed to get sufficient sleep, keep fit, eat a sufficient breakfast, and consume their normal amounts of caffeinated food and drink on the day of cognitive testing and MRI scanning. In addition, subjects were instructed to avoid alcohol the night before the assessment.

The descriptions in this subsection are mostly reproduced from another study within the same project that used the exact same methods (Takeuchi et al., 2015a).

Details of some psychological measures

Tanaka B-type intelligence test (TBIT) type 3B.

Tanaka B-type intelligence test (Tanaka et al., 2003). Type 3B, which is for examinees in

their 3rd-year of junior high school and older, was used in this study. This test is a nonverbal mass intelligence test which does not include story problems but uses figures, single numbers, and letters as stimuli. In all subtests, subjects had to complete as many problems as possible within a certain time (a few minutes). This test consist of a maze test (subjects had to trace a maze with a pencil from start to finish), counting cubes (subjects had to count the number of cubes piled up in three-dimensional ways), a displacement task (figures and numbers; subjects had to substitute a figure [9 figures] with a number [1 to 9] according to a model chart), identification vs. same-different judgments (Japanese kana characters; subjects had to judge whether a pair of meaningless Japanese strings were the same), filling in a sequence of numbers (subjects had to fill in the blanks of a number sequence with suitable numbers according to the rules of the number arrangement), marking figures (subjects had to select forms which were identical to three samples from a series [sequence] of eight different forms), and filling in figures (subjects had to complete uncompleted figures so that the uncompleted figures were the same as the sample figures when rotated).

The perception speed factor of TBIT measures simple processing speed. This factor involves three subtests: a displacement task, identification vs. same–different judgments, and marking figures.

These descriptions were mostly reproduced from our previous study (Takeuchi et al., 2011a).

A (computerized) digit span task.

Computerized forward and backward digit span tests were used to assess verbal working memory performance. In this task, subjects were asked to view a progressively

increasing number of random digits visually presented one-digit per second on a computer screen. They were then asked to repeat the sequence by pressing numbered buttons on the screen in the presented order (digit-span forward) or in the reverse order (digit-span backward), starting from two digits. Three sequences were given at each level, until the participants responded incorrectly to all three sequences, at which point the task was ended. The score of each test is equal to the sum of the number of digits correctly repeated in the digit span forward and digit span backward tasks.
These descriptions were mostly reproduced from our previous study (Takeuchi et al., 2011a).

Systemizing Quotient (SQ) and Empathy Quotient (EQ) questionnaires.

Japanese versions (Wakabayashi et al., 2007) of the SQ and EQ questionnaires (Baron-Cohen et al., 2003; Baron-Cohen and Wheelwright, 2004) were administered. The following methods were reproduced from our previous study using the exact same method (Takeuchi et al., 2014a; Takeuchi et al., 2014b; Takeuchi et al., 2013). The EQ score was used as an index of empathizing, and the SQ score was used as an index of systemizing. These tests consist of 40 items for each quotient and 20 filler items that are not scored. The scales consist of self-descriptive statements scored on a four-point scale ranging from Strongly Disagree to Strongly Agree. Half the items are worded to produce an ‘‘agree’’ response and rest to produce a ‘‘disagree’’ response. Items are

randomized to avoid a response bias. Each strong systemizing/empathizing response is awarded 2 points, and each slightly systemizing/empathizing response is awarded 1 point (i.e., each item is scored as 2, 1, or 0), resulting in a range of total scores from 0–80 for each quotient.

The questionnaire comprised the psychometric properties described below. Some studies have reported empathizing and systemizing as largely independent. However, a weak negative correlation between them has been reported by several studies(e.g., Wheelwright et al., 2006), whereas others failed to find such correlation(e.g., Wakabayashi et al., 2007). Individuals with autism spectrum conditions have been found to exhibit higher SQ scores and lower EQ scores than controls (Wakabayashi et al., 2007). Similarly, male individuals exhibit higher SQ scores than female individuals who, in turn, present higher EQ scores than males (Wakabayashi et al., 2006b). Students of humanities also show higher EQ scores than students of science who, in turn, present higher SQ scores than those studying humanities (Wakabayashi et al., 2006b). Additionally, actors were found to have higher EQ scores (Nettle, 2006). EQ is positively correlated with both the size of an individual's social network (Stileman, 2007) and their performance on a face perception task (Penton-Voak et al., 2007). The Autism Spectrum Quotient (AQ) is a measure of autistic traits. Although that measure was not collected in this project, it is well explained by the model including both EQ and SQ (more than 75% of the variance) (Wheelwright et al., 2006) (Groen et al., 2015). These findings have demonstrated the criterion-related validity of the present questionnaire. The internal consistencies of both EQ and SQ, calculated in a previous,

large sample study, were 0.86 and 0.88, respectively, demonstrating the reliability of this questionnaire.

The Japanese version of the questionnaires was validated by Prof. Akio Wakabayashi, Prof. Baron-Cohen, and others (Wakabayashi et al., 2007; Wakabayashi et al., 2006a). In the Japanese version, the patterns of male EQ and SQ scores (vs. female EQ and SQ scores), ASC group (vs. controls), and science majors (vs. humanities majors) were similar to those of the original version (Wakabayashi et al., 2007; Wakabayashi et al., 2006a). The present study's participants' EQ score was lower than that of the previous study's control sample (Baron-Cohen et al., 2003; Baron-Cohen and Wheelwright, 2004), but similar to the EQ scores shown by the previous study's Japanese university students (Wakabayashi et al., 2007; Wakabayashi et al., 2006a).

The following are examples of items found on the SQ–EQ questionnaires:

"I can tune into how someone else feels rapidly and intuitively" (EQ)

"I am good at predicting how someone will feel" (EQ)

"I am fascinated by how machines work" (SQ)

"If I were buying a stereo, I would want to know about its precise technical features" (SQ)

These descriptions were mostly reproduced from our previous study (Takeuchi et al., 2019).

Preprocessing of imaging data

Preprocessing and analysis of functional activation data were performed using SPM8 implemented in Matlab. The following descriptions are reproduced from our previous study using the same preprocessing procedure (Takeuchi et al., 2015b). Before analysis, individual BOLD images were re-aligned and resliced to the mean BOLD image and they were then corrected for slice timing. And the abovementioned mean BOLD image was then realigned to the mean b = 0 image as previously described together with slice timing corrected images *(Takeuchi et al., 2011b).* As the mean b = 0 image was aligned with the FA image and MD map, the BOLD image, b = 0 image, FA image, and MD map were all aligned. For normalization, we used a previously validated two-step "new segmentation" algorithm of diffusion images and a previously validated diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL)-based registration process (Takeuchi et al., 2013). This normalization method was used for all diffusion images, including gray matter segments (regional gray matter density [rGMD] map), white matter segments (regional white matter density [rWMD] map), and cerebrospinal fluid (CSF) segments (regional CSF density [rCSFD] map).

Using the new segmentation algorithm implemented in SPM8, FA images from each individual were segmented into six tissues (first new segmentation). In this process, default parameters and tissue probability maps were used, except that affine regularization was performed using the International Consortium for Brain Mapping (ICBM) template for East Asian brains, and the sampling distance (the approximate distance between sampled points when estimating the model parameters) was 2 mm. We next synthesized the FA image and MD map. In this synthesized image, the area with WM tissue probability > 0.5 in the aforementioned new segmentation process was the FA image multiplied by −1. Hence, this synthesized image shows a very clear contrast between WM and other tissues. The remaining area is the MD map. The reasons for this procedure are given below. The synthesized image from each individual was then segmented using the new segmentation algorithm implemented in SPM8 and the same parameters as above (second new segmentation). This two-step segmentation process was adopted because the FA image had a relatively clear contrast between GM and WM, as well as between WM and CSF, so the first new segmentation step can distinguish WM from the other tissues. On the other hand, the MD map had a clear contrast between GM and CSF, and the second new segmentation can segment GM. Since the MD map alone

does not have a clear contrast between WM and GM, we needed to use a synthesized image and the two-step segmentation process.

We next proceeded to the DARTEL registration process implemented in SPM8. In this process, we used the DARTEL import image of the GM tissue probability map produced by the second new segmentation process as the GM input for the DARTEL process. The WM input for the DARTEL process was created as follows. First, the raw FA image was multiplied by the WM tissue probability map from the second new segmentation process within the areas with a WM probability > 0.5 (the signals from all other areas were set to 0). Then, this FA image × the WM tissue probability map was coregistered and resliced to the DARTEL import WM tissue probability image from the second segmentation. The template for the DARTEL procedures was created using imaging data from 63 subjects who participated in (Takeuchi et al., 2011b) and in the present study. The use of a subset of subjects to produce the template was necessary because the project is still ongoing, and we cannot reprocess all images using a new template every time we add subjects, especially as our DARTEL processes require a great deal of time and the resulting images require substantial storage resources. However, this sample (N = 63) is not small compared to previous studies, so we believe that these subjects have the same characteristics on average as the entire cohort.

Next, using this existing template, DARTEL procedures were performed. The parameters were changed as follows to improve the accuracy of the procedures. The number of Gauss–Newton iterations to be performed within each outer iteration was set to 10. In each outer iteration, we used 8-fold more timepoints than the default values to solve the partial differential equations. The number of cycles used by the full multi-grid matrix solver was set to 8. The number of relaxation iterations performed in each multi-grid cycle was also set to 8. The resultant synthesized images were spatially normalized to Montreal Neurological Institute (MNI) space. Using the parameters for these procedures, the raw FA, rGMD, and rWMD maps from the abovementioned second new segmentation process were normalized to yield images with 1.5 × 1.5 × 1.5 mm3 voxels. The FA image * the WM tissue probability map was used in the DARTEL procedures because this synthesized image includes different signal intensities within the WM tissues, and the normalization procedure can take advantage of these intensity differences to adjust the image to the template from the outer edge of the tissue and within the WM tissue. No modulation was performed in this normalization procedure.

The voxel size of normalized FA images, MD images, and segmented images was 1.5 × 1.5 × 1.5 mm3. The voxel size of normalized BOLD images was 3 × 3 × 3

mm3. Next, we created averaged normalized rGMD and rWMD images using a subset of the entire sample (63 subjects) (Takeuchi et al., 2013).

We did not co-register EPI or DTI images to T1-weighted structural images because they had a different shape resulting from the unignorable distortion of EPI images acquired by 3T MRI.

Rationale for including the RAPM score as a covariate in whole-brain imaging analyses

We initially did not add RAPM score as a covariate in part because psychometric intelligence and closely related measures are the dependent variables in psychological analyses. Adding psychometric intelligence measure also creates analytic problems, as psychometric intelligence is supposed to be a fundamental psychological correlate of originality (relative to fluency) as described in the Introduction of the main text.

On the contrary, it is common to add a measure of general intelligence as a covariate in neuroimaging studies of individual differences in cognition. Thus, we added the RAPM as a covariable in the whole brain analyses. However, the significance of the clusters presented in the main text remained significant after removing RAPM score from the covariate set and this procedure does not affect the conclusion and discussions of the main text.

Supplemental Results

Psychological main effects and interactions of total score

ANCOVA revealed significant main effects of the total score of S-A creativity test on the total score of TBIT (an intelligence test consisting of speeded tasks), perception factor score on TBIT, vigor subscale on POMS, EQ, and SQ. These findings are consistent with work by our group and other researchers (Jauk et al., 2013; Preckel et al., 2011; Takeuchi et al., 2016; Takeuchi et al., 2014c).

As for the interactions, ANCOVA revealed a significant interaction between sex and total score on the S-A creativity test with EQ score (stronger positive correlation in males than females). See Supplemental Table 2 for full results.

Main effects of total score on the S-A creativity test on functional activation

ANCOVA revealed an overall positive main effect (regardless of sex) of the total score on the S-A creativity test on functional activity during the 0-back task in a cluster mainly around the right supramarginal gyrus, right superior and middle temporal gyrus, and right inferior frontal gyrus (most (86.9%) of this cluster belongs to areas activated during the 0-back task), a cluster that spread from the right inferior frontal orbital area to the right insula (all of this cluster belongs to areas deactivated during the 0-back task), a cluster that spread from the right inferior frontal operculum to the right inferior frontal triangular (all of this cluster belongs to areas activated during the 0-back

task), and a cluster that spread from other middle frontal areas to the superior frontal gyrus (located at the border between areas activated during the 0-back task and areas deactivated, including both of areas) (Supplemental Fig. 1a). For full statistical results, see Supplemental Table 3.

In addition, ANCOVA revealed an overall positive main effect (regardless of sex) of total score on the S-A creativity test on functional activity during the 2-back task in a cluster mainly around the right supramarginal gyrus, right angular gyrus, right postcentral gyrus, right superior, middle and inferior temporal gyrus, right temporal pole, right inferior frontal gyrus, right fusiform gyrus, and right parahippocampal gyrus, (47.8% and 40.3% of this large cluster belongs to areas activated and deactivated during the 2-back task, respectively); two anatomical clusters mostly spread around the right inferior and middle frontal orbital area, inferior frontal triangular, and right insula (most—86.9% and 100%—of these clusters belong to areas deactivated during the 2-back task]; an anatomical cluster spread around the right fusiform gyrus, right parahippocampal gyrus, and right cerebellum (all voxels of this cluster belong to areas deactivated during the 2-back task); an anatomical cluster spread around the left middle and superior temporal gyrus; an anatomical cluster spread around the right thalamus (all voxels in this and the previously described cluster belong to areas activated during the

2-back task); an anatomical cluster spread around the left middle and superior temporal gyrus (this area is located at the border between areas activated during the 2-back task and those deactivated, including both areas); and an anatomical cluster spread around the right precuneus and right cuneus (most (80%) of this cluster belongs to areas deactivated during the 2-back task) (Supplemental Fig. 1b). The findings for the last cluster are consistent with our previous work (Takeuchi et al., 2011b). For all statistical results, see Supplemental Table 3.

Interaction effect of sex and total scores on functional activation

ANCOVA revealed an interaction effect between sex and total score on the S-A creativity test on functional activity during the 0-back task in a cluster spread mainly around the left supramarginal gyrus, left inferior parietal lobule, and left postcentral gyrus (all voxels in this cluster belong to areas activated during the 0-back task), and in clusters spread around the left precentral gyrus (most voxels in these clusters belong to areas activated during the 0-back task). These interactions were formed with the positive correlation in males and negative correlations in females (Supplemental Fig. 2).

For full statistical results, see Supplemental Table 4.

Supplemental Discussion

Discussions on the difference in neural correlates between originality/fluency score and total CMDT score

While a number of commonalities of correlates were observed between originality/fluency score and total CMDT score, there were also certain differences in the neural correlates between originality/fluency score and total CMDT score, even though originality per idea (originality/fluency) contributes to total CMDT score. Greater originality/fluency showed significant widespread associations with lower deactivation in the DMN in response to the contrast of 2-back and 0-back, whereas total CMDT scores did not show significant associations with brain activation in relation to this contrast. As discussed in the main text, these findings are partly common in the association between general intelligence measured by non-verbal reasoning and brain activity-related to the 2-back and 0-back contrast (Takeuchi et al., 2018). Additionally, these results may be in parallel with the lack of correlation between general intelligence measured by non-verbal reasoning and the CMDT total score, whereas originality/fluency scores showed a significant positive correlation that was observed in this study. And these contrasts of the correlates may suggest that intelligence and relevant neural mechanisms are more important for generating original ideas compared with total CMDT scores.

Discussions in the difference on the findings of the association between brain activity during the 2-back WM task and total CMDT score between the present study and our previous study

Our previous study revealed that higher total CMDT score was associated with lower reduced task-induced deactivation during the 2-back WM task in the posterior medial parietal cortex (Takeuchi et al., 2011b). On the other hand, in the present study, the same association patterns were seen in the overlapping area of the posterior cingulate cortex (Supplemental Fig. 2). These significant associations were most prominently observed in the extensive areas of the ventral attention system (Corbetta et al., 2008), which spread across the areas of activation and deactivation during tasks. Statistically, the difference may be partly attributed to the greater statistical power or much larger sample size in this study compared with the previous study (N = 1221 v.s. N = 63) and the former may be able to accurately depict the sum of associations between total CMDT scores and attention-related brain activity. Furthermore, in the present study, we used TFCE-based permutation tests for correcting multiple comparisons, whereas in the our previous study, we used a cluster size test that was based on the random field theory and used the lenient voxel-level cluster determining threshold, which is difficult to use for substantially localizing significant findings.

Functionally, as described in the main text, the present finding may suggest that task-induced activation in the DMN and the activation of the ventral attention network during cognitive tasks, which may reflect flexible attention shifts (Petersen and Posner, 2012), are important for total CMDT score This result is consistent with a recent review (Zabelina, 2018).

Low correlation levels between brain activity and CMDT scores

The simple correlation coefficients between beta estimates of functional activity and creativity measured by divergent thinking (CMDT) scores were relatively small (< |0.2| in each sex) in the present study. As also discussed in our previous study (Takeuchi et al., 2018), relatively weak correlations ($r < 0.2$) between individual cognitive differences and neuroimaging measures are universal phenomenon in studies with massive samples of young normal adults (i.e., N > several hundred) regardless of the type of imaging measures (Magistro et al., 2015; Schilling et al., 2012; Takeuchi et al., 2015a; Takeuchi et al., 2018; Takeuchi et al., 2017) without exception to our knowledge. Therefore, the low correlations are not indicative of the low importance of our findings. It should be noted that the large correlation coefficients for peak voxels in significant areas in the whole-brain analyses with small sample sizes were extremely

overestimated due to strong overfitting effects and are not reflective of true effect size (e.g., Murphy et al., 2012; Vul et al., 2009).

Potential limitations

This study is subject to several limitations. One limitation common to our previous studies, as well as many other studies in this field, is the use of college student cohorts (Jung et al., 2010; Song et al., 2008; Takeuchi et al., 2011c; Takeuchi et al., 2010a, b; Wei et al., 2013). Whether findings from young, educated samples hold across all education levels and age ranges should be confirmed in future studies.

References

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., Wheelwright, S., 2003. The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358, 361-374.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. Journal of Autism and Developmental Disorders 34, 163-175.

Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain: from environment to theory of mind. Neuron 58, 306-324.

Groen, Y., Fuermaier, A., Den Heijer, A., Tucha, O., Althaus, M., 2015. The empathy and systemizing quotient: The psychometric properties of the Dutch version and a review of the cross-cultural stability. Journal of Autism and Developmental Disorders 45, 2848-2864.

Jauk, E., Benedek, M., Dunst, B., Neubauer, A.C., 2013. The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. Intelligence 41, 212-221.

Jung, R.E., Segall, J.M., Bockholt, H.J., Flores, R.A., Smith, S.M., Chavez, R.S., Haier, R.J., 2010. Neuroanatomy of creativity. Human brain mapping 31, 398-409.

Magistro, D., Takeuchi, H., Nejad, K.K., Taki, Y., Sekiguchi, A., Nouchi, R., Kotozaki, Y., Nakagawa, S., Miyauchi, C.M., Iizuka, K., Yokoyama, R., Shinada, T., Yamamoto, Y., Hanawa, S., Araki, T., Hashizume, H., Sassa, Y., Kawashima, R., 2015. The Relationship between Processing Speed and Regional White Matter Volume in Healthy Young People. PLoS ONE 10, e0136386.

Maldjian, J.A., Laurienti, P.J., Burdette, J.H., 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. Neuroimage 21, 450-455.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233-1239.

Murphy, S., Norbury, R., Godlewska, B., Cowen, P., Mannie, Z., Harmer, C., Munafo, M., 2012. The effect of the serotonin transporter polymorphism (5-HTTLPR) on amygdala function: a meta-analysis. Molecular Psychiatry 18, 512-520.

Nettle, D., 2006. Psychological profiles of professional actors. Personality and individual differences 40, 375-383.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97-113.

Penton-Voak, I.S., Allen, T., Morrison, E.R., Gralewski, L., Campbell, N., 2007. Performance on a face perception task is associated with empathy quotient scores, but not systemizing scores or participant sex. Personality and individual differences 43, 2229-2236.

Petersen, S.E., Posner, M.I., 2012. The attention system of the human brain: 20 years after. Annual Review of Neuroscience 35, 73-89.

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

Preckel, F., Wermer, C., Spinath, F.M., 2011. The interrelationship between speeded and unspeeded divergent thinking and reasoning, and the role of mental speed. Intelligence 39, 378-388.

Schilling, C., Kühn, S., Paus, T., Romanowski, A., Banaschewski, T., Barbot, A., Barker, G., Brühl, R., Büchel, C., Conrod, P., 2012. Cortical thickness of superior frontal cortex predicts impulsiveness and perceptual reasoning in adolescence. Molecular Psychiatry 18, 624-630.

Song, M., Zhou, Y., Li, J., Liu, Y., Tian, L., Yu, C., Jiang, T., 2008. Brain spontaneous functional connectivity and intelligence. Neuroimage 41, 1168-1176.

Stileman, E., 2007. Construction of the Social Network Score (SNS) Questionnaire for undergraduate students, and an examination of the pre-requisites for large social networks in humans? Undergraduate thesis., Department of Psychology. University of Edinburgh, Edinburgh.

Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., Kawashima, R., 2011a. Effects of training of processing speed on neural systems. Journal of Neuroscience 31, 12139-12148.

Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., Kawashima, R., 2011b. Failing to deactivate: the association between brain activity during a working memory task and creativity. Neuroimage 55, 681-687.

Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., Kawashima, R., 2011c. Failing to deactivate: the association between brain activity during a working memory task and creativity. Neuroimage 55, 681-687.

Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Hashizume, H., Sassa, Y., Kotozaki, Y., Miyauchi, C.M., Yokoyama, R., Iizuka, K., 2014a. Association between resting-state functional connectivity and empathizing/systemizing. Neuroimage 99, 312-322.

Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Hashizume, H., Sassa, Y., Kotozaki, Y., Miyauchi, C.M., Yokoyama, R., Iizuka, K., Seishu, N., Tomomi, N., Kunitoki, K., Kawashima, R., 2015a. Degree centrality and fractional amplitude of low-frequency oscillations associated with Stroop interference. Neuroimage 119, 197-209.

Takeuchi, H., Taki, Y., Nouchi, R., Yokoyama, R., Kotozaki, Y., Nakagawa, S., Sekiguchi, A., Iizuka, K., Hanawa, S., Araki, T., Miyauchi, C.M., Sakaki, K., Sassa, Y., Nozawa, T., Ikeda, S., Yokota, S., Daniele, M., Kawashima, R., 2018. General intelligence is associated with working memory-related brain activity: new evidence from a large sample study. Brain Structure and Function Epub ahead of print.

Takeuchi, H., Taki, Y., Nouchi, R., Yokoyama, R., Kotozaki, Y., Nakagawa, S., Sekiguchi, A., Iizuka, K., Yamamoto, Y., Hanawa, S., Araki, T., Miyauchi, C.M., Sakaki, K., Sassa, Y., Nozawa, T., Ikeda, S., Yokota, S., Magistro, D., Kawashima, R., 2019. Empathizing associates with mean diffusivity. Scientific reports 9, Article number: 8856.

Takeuchi, H., Taki, Y., Nouchi, R., Yokoyama, R., Kotozaki, Y., Nakagawa, S., Sekiguchi, A., Iizuka, K., Yamamoto, Y., Hanawa, S., Araki, T., Miyauchi, M., Calros , Shinada, T., Sakaki, K., Sassa, Y., Nozawa, T., Ikeda, S., Yokota, S., Daniele, M., Kawashima, R., 2017. Global associations between regional gray matter volume and diverse complex cognitive functions: evidence from a large sample study. Scientific reports 7, article 10014.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., Kawashima, R., 2010a. Regional gray matter volume of dopaminergic system associate with creativity: Evidence from voxel-based morphometry Neuroimage 51, 578-585.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., Kawashima, R., 2010b. White matter structures associated with creativity: Evidence from diffusion tensor imaging. Neuroimage 51, 11-18.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., Kawashima, R., 2014b. Regional gray matter volume is associated with empathizing and systemizing in young adults. PLoS ONE 9, e84782.

Takeuchi, H., Taki, Y., Sekiguchi, A., Nouchi, R., Kotozaki, Y., Nakagawa, S., Miyauchi, C.M., Iizuka, K., Yokoyama, R., Shinada, T., 2016. Mean diffusivity of basal ganglia and thalamus specifically associated with motivational states among mood states. Brain Structure and Function, 1-11.

Takeuchi, H., Taki, Y., Sekiguchi, A., Nouchi, R., Kotozaki, Y., Nakagawa, S., Miyauchi, C.M., Iizuka, K., Yokoyama, R., Shinada, T., Yamamoto, Y., Hanawa, S., Araki, T., Hashizume, H., Sassa, Y., Kawashima, R., 2014c. Creativity measured by divergent thinking is associated with two axes of autistic characteristics. Frontiers in psychology 5, Article 921, 1-8.

Takeuchi, H., Taki, Y., Thyreau, B., Sassa, Y., Hashizume, H., Sekiguchi, A., Nagase, T., Nouchi, R., Fukushima, A., Kawashima, R., 2013. White matter structures associated with empathizing and systemizing in young adults. Neuroimage 77, 222-236.

Takeuchi, H., Tomita, H., Taki, Y., Kikuchi, Y., Ono, C., Yu, Z., Sekiguchi, A., Nouchi, R., Kotozaki, Y., Nakagawa, S., 2015b. Cognitive and neural correlates of the 5-repeat allele of the dopamine D4 receptor gene in a population lacking the 7-repeat allele. Neuroimage 110, 124-135.

Tanaka, K., Okamoto, K., Tanaka, H., 2003. Manual of New Tanaka B type intelligence test. Kaneko Syobo, Tokyo.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273-289.

Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Reply to comments on "puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition". Perspectives on Psychological Science 4, 319-324.

Wakabayashi, A., Baron-Cohen, S., Uchiyama, T., Yoshida, Y., Kuroda, M., Wheelwright, S., 2007. Empathizing and systemizing in adults with and without autism spectrum conditions: cross-cultural stability. Journal of Autism and Developmental Disorders 37, 1823-1832.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., 2006a. Individual and gender differences in Empathizing and Systemizing: Measurement of individual differences by the Empathy Quotient (EQ) and the Systemizing Quotient (SQ). Japanese Journal of Psychology 77, 271-277.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., Weil, L., 2006b. Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). Personality and individual differences 41, 929-940.

Wei, D., Yang, J., Li, W., Wang, K., Zhang, Q., Qiu, J., 2013. Increased resting functional connectivity of the medial prefrontal cortex in creativity by means of cognitive stimulation. Cortex 51, 92-102.

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., Weil, L., Wakabayashi, A., 2006. Predicting autism spectrum quotient (AQ) from the

systemizing quotient-revised (SQ-R) and empathy quotient (EQ). Brain Research 1079, 47-56.

Zabelina, D.L., 2018. Attention and creativity. In: Jung, R., Vartanian, O. (Eds.), The Cambridge handbook of the neuroscience of creativity. Cambridge University Press, Cambridge, MA, pp. 161-179.

Supplemental Table 1.

Demographics of study participants.

	Male		Female	
Measure	Mean	SD	Mean	SD
Age	20.81	1.89	20.67	1.62
S-A creativity-total	37.56	10.64	39.23	9.96
S-A creativity-originality	8.37	3.32	8.23	3.24
S-A creativity-fluency	34.44	9.15	35.4	8.46
S-A creativity-flexibility	24.93	5.36	25.76	4.89
S-A creativity-elaboration	29.19	8.3	30.99	7.8
Originality/fluency	0.2438	0.0815	0.2326	0.0784
RAPM	28.8	3.8	27.98	3.84
0-back-RT (ms)	4521.19	753.6	4569.88	777.8
0-back-accuracy (%)	1.00	0.02	1.00	0.02
2-back-RT (ms)	6712.12	1771.52	6754.45	1821.69
2-back-accuracy (%)	0.99	0.08	1.00	0.01
Volume-wise frame-wise Displacement*	0.20	0.05	0.21	0.05

*Frame-wise displacement was calculated using the method by Power et al. (2012)

Supplemental Table 2. Main and interaction effects of ANCOVAs for total CMDT scores on psychological measures.

	Total score			
	Main effect [F score, P value (unc), P value (FDR)]	Interaction [F score, P value (unc), P value (FDR)]	Male correlation (r)	Female correlation (r)
RAPM	2.95	0.40	0.031	0.063
(M:700, F:521)	0.086	0.526		
	0.120	0.460		
TBIT-Total score	52.79	2.58	0.175	0.267
(M:635, F:468)	7.03×10^{-13}	0.108		
	$1.48 \times 10^{-11*}$	0.142		
TBIT-Perception	26.86	3.737	0.100	0.221
speed factor	2.61×10^{-7}	0.053		
(M:635, F:468)	$1.83 \times 10^{-6*}$	0.086		

Digit span	2.17	0.26	0.058	0.030
(M:635, F:468)	0.141	0.612		
	0.174	0.460		
Vigor scale of	37.40	0.30	0.202	0.171
POMS	1.32×10^{-9}	0.585		
(M:656, F:486)	$1.39 \times 10^{-8*}$	0.460		
Empathizing	16.88	5.08	0.198	0.065
(M:700, F:521)	4.25×10^{-5}	0.024		
	$1.79 \times 10^{-4*}$	0.046*		
Systemizing	17.30	0.924	0.148	0.106
(M:700, F:521)	3.42×10^{-5}	0.337		
	$1.79 \times 10^{-4*}$	0.337		

Supplemental Table 3. Brain regions showing significant main effects for total S-A creativity test score on brain activity.

Including gray matter areas*	x	y	z	TFCE score	Corrected P value (FWE, TFCE)	Cluster size (voxels)	activated/ deactivated **	r*** (male/ female)
Positive main effect of total score on activity during the 0-back task								
Supramarginal gyrus (R)/ Middle temporal gyrus (R)/ Superior temporal gyrus (R)/	66	-33	-3	640.13	0.0142	873	759/49	0.171/0.079
Inferior frontal orbital area (R)/Insula (R)/	30	24	-15	471.86	0.0448	10	0/7	0.071/0.184
Inferior frontal operculum (R)/Inferior frontal triangular (R)/	57	12	9	465.66	0.0464	24	21/0	0.090/0.128
Middle frontal other areas (R)/Superior frontal other areas (R)/	33	54	12	460.48	0.0474	23	7/10	0.093/0.135
Positive main effect of total score on activity during the 2-back task								
Supramarginal gyrus (R)/Inferior temporal gyrus (R)/Middle temporal gyrus (R)/Temporal pole (R)/Superior temporal gyrus (R)/	60	-42	9	982.68	0.0008	2179	1042/878	0.122/0.149
Inferior frontal orbital area (R)/Inferior frontal triangular (R)/	51	30	-3	520.56	0.02	160	13/139	0.063/0.172
Inferior frontal orbital area (R)/Middle frontal orbital area (R)/Insula (R)/	33	27	-18	451.1	0.0352	44	0/42	0.054/0.183
Fusiform gyrus (R)/Parahippocampal gyrus (R)/Cerebellum (R)/	27	-30	-24	432.33	0.042	4	0/4	0.033/0.107
Middle temporal gyrus (L)/Superior temporal gyrus (L)/	-66	-42	9	429.29	0.0426	22	20/0	0.118/0.104
Thalamus (R)/	9	-12	0	422.77	0.0464	8	8/0	0.059/0.126
Middle temporal gyrus (L)/Superior temporal gyrus (L)/	-45	-39	6	422.66	0.0464	9	4/3	0.130/0.070
Cuneus (R)/Precuneus (R)/	18	-66	27	418.27	0.0492	5	1/4	0.066/0.129

*Labeling of the anatomical regions of gray matter is based on the WFU PickAtlas Tool (http://www.fmri.wfubmc.edu/cms/software#PickAtlas/) (Maldjian et al., 2004; Maldjian et al., 2003) and the PickAtlas automated anatomical labeling atlas option (Tzourio-Mazoyer et al., 2002). In this atlas, temporal pole areas and some other areas include all subregions. Areas of the superior frontal other areas include areas of the superior frontal gyrus other than the medial, orbital, and medial-orbital parts of the superior frontal gyrus. Only areas with significant voxels comprising 10% or more of the cluster or areas with 50 or more significant voxels that existed in the cluster were reported.

**Percentage of voxels activated or deactivated during the corresponding condition (i.e., in the case of 0-back analyses, activity/deactivation during the 0-back task) among the whole sample ($P < 0.05$, FDR corrected at the voxel level).

*** Simple correlation coefficients between mean beta estimates of significant clusters and psychological scores. Note that due to overfitting in whole-brain analyses (Vul et al., 2009), the correlation coefficients of significant areas are overestimated to a degree depending on the sample size and number of comparisons.

Supplemental Table 4. Brain regions showing significant interactions between total scores and sex on brain activity.

Including gray matter areas*	x	y	z	TFCE score	Corrected P value (FWE, TFCE)	Cluster size (voxels)	activated/ deactivated **	r*** (male/ female)
Interaction between total score and sex on 0-back activity (positive effects in males and negative effects in females)								
Inferior parietal lobule (L)/Postcentral gyrus (L)/	-36	-30	45	495.59	0.0412	51	51/0	0.140/-0.138
Precentral gyrus (L)/	-30	-15	60	476.33	0.0464	9	9/0	0.113/-0.120
Precentral gyrus (L)/	-30	-18	42	470.58	0.0484	8	7/1	0.112/-0.125

*Labeling of the anatomical regions of gray matter is based on the WFU PickAtlas Tool (http://www.fmri.wfubmc.edu/cms/software#PickAtlas/) (Maldjian et al., 2004; Maldjian et al., 2003) and the PickAtlas automated anatomical labeling atlas option (Tzourio-Mazoyer et al., 2002). In this atlas, temporal pole areas and some other areas include all subregions. Areas of the superior frontal other areas include areas of the superior frontal gyrus other than the medial, orbital, and medial-orbital parts of the superior frontal gyrus. Only areas with significant voxels comprising 10% or more of the cluster or areas with 50 or more significant voxels that existed in the cluster were reported.

**Percentage of voxels activated or deactivated during the corresponding condition (i.e., in the case of 0-back analyses, activity/deactivation during the 0-back task) among the whole sample ($P < 0.05$, FDR corrected at the voxel level).

*** Simple correlation coefficients between mean beta estimates of significant clusters and psychological scores. Note that due to overfitting in whole-brain analyses (Vul et al., 2009), the correlation coefficients of significant areas are overestimated to a degree depending on the sample size and number of comparisons.

Supplemental Figure legends

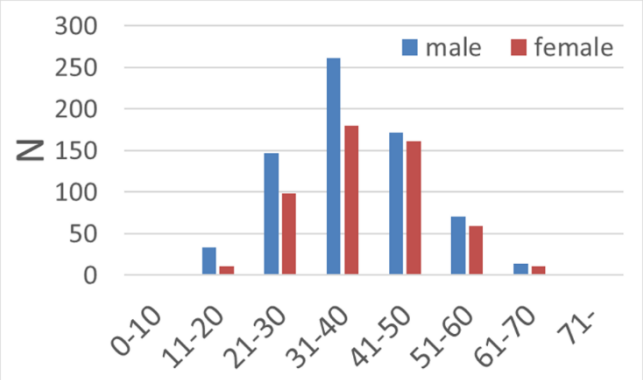
Supplemental Fig. 1. Distribution of the total scores from the S-A creativity test in our sample.

Supplemental Fig. 2. Regions with a significant positive main effect of total CMDT score on brain activity. (Left panels) Total score on the S-A creativity test showed a significant positive main effect on (a) brain activity during the 0-back task , and (b) that during the 2-back task. The results were obtained using a threshold of $P < 0.05$, corrected for multiple comparisons based on 5000 permutations using TFCE scores. (Middle panels) Scatterplots of the associations between total score on the S-A creativity test and mean beta estimates of significant clusters. (Right panels) Areas deactivated during the corresponding conditions. All results are displayed at a height threshold of 0.05, FDR corrected. (Left and right panels) Results rendered on a "render" image or a "single-subject T1" image (in the case of section images) in SPM8.

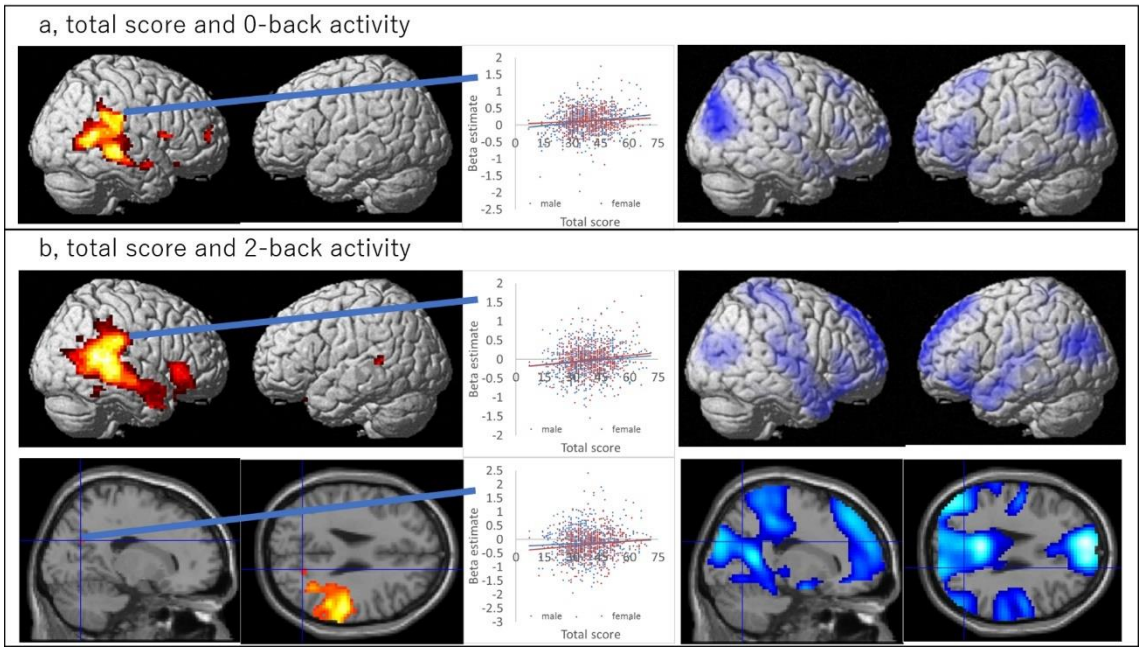
Supplemental Fig. 3. Regions with significant interaction effects between sex and scores on the CMDT on brain activity. (Left panels) Total score on the S-A creativity test showed significant interaction effects with sex on brain activity during the 0-back task. This interaction was moderated by a positive correlation in males and negative correlation in females. Results were obtained using a threshold of $P < 0.05$, corrected for multiple comparisons based on 5000 permutations using TFCE scores. (Middle panels) Scatterplots of the associations between scores on the S-A creativity test and

mean beta estimates of significant clusters. (Right panels) Areas activated (red) and deactivated (blue) during the corresponding conditions. All results are displayed at a height threshold of 0.05, FDR corrected. (Left and right panels) Results are rendered on a “render” image or a “single-subject T1” image (in the case of section images) in SPM8.

Supplemental Figure 1.



Supplemental Figure 2.



Supplemental Figure 3.