

General intelligence is associated with working memory-related brain activity: new evidence from a large sample study

Hikaru Takeuchi¹ · Yasuyuki Taki^{1,2,3} · Rui Nouchi^{4,5,6} · Ryoichi Yokoyama⁷ · Yuka Kotozaki⁸ · Seishu Nakagawa^{9,10} · Atsushi Sekiguchi^{2,11} · Kunio Iizuka¹² · Sugiko Hanawa⁹ · Tsuyoshi Araki⁶ · Carlos Makoto Miyauchi¹³ · Kohei Sakaki¹⁰ · Yuko Sassa¹ · Takayuki Nozawa¹⁴ · Shigeyuki Ikeda¹⁵ · Susumu Yokota¹ · Magistro Daniele¹⁶ · Ryuta Kawashima^{1,6,14}

¹ *Division of Developmental Cognitive Neuroscience, Institute of Development, Aging and Cancer (IDAC), Tohoku University, 4-1 Seiryō-cho, Aoba-ku, Sendai 980-8575, Japan*

² *Division of Medical Neuroimaging Analysis, Department of Community Medical Supports, Tohoku Medical Megabank Organization, Tohoku University, Sendai, Japan*

³ *Department of Radiology and Nuclear Medicine, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan*

⁴ *Creative Interdisciplinary Research Division, Frontier Research Institute for Interdisciplinary Science, Tohoku University, Sendai, Japan*

⁵ *Human and Social Response Research Division, International Research Institute of Disaster Science, Tohoku University, Sendai, Japan*

⁶ *Advanced Brain Science, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan*

⁷ *School of Medicine, Kobe University, Kobe, Japan*

⁸*Division of Clinical research, Medical-Industry Translational Research Center, Fukushima Medical University School of Medicine, Fukushima, Japan*

⁹*Department of Functional Human Brain Imaging Science, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan*

¹⁰*Division of Psychiatry, Tohoku Medical and Pharmaceutical University, Sendai, Japan*

¹¹*Department of Psychosomatic Research, National Institute of Mental Health, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan*

¹²*Department of Psychiatry, Tohoku University Graduate School of Medicine, Sendai, Japan*

¹³*Department of Language Sciences, Graduate School of Humanities, Tokyo Metropolitan University, Tokyo, Japan.*

¹⁴*Collaborative Research Center for Happiness Co-creation Society through "Ishin-Denshin" Intelligent Communications, Tokyo Institute of Technology, Tokyo, Japan*

¹⁵*Department of Ubiquitous Sensing, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan*

¹⁶*Department of Sport Science, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom*

Corresponding author:

Hikaru Takeuchi

Division of Developmental Cognitive Neuroscience, IDAC, Tohoku University

4-1 Seiryō-cho, Aoba-ku, Sendai 980-8575, Japan

Tel/Fax: +81-22-717-7988

E-mail: takehi@idac.tohoku.ac.jp

Keywords:

General intelligence; Non-verbal reasoning; Working memory; Brain activity; N-back task

Short title: Intelligence associated with brain activity

Abstract

Psychometric intelligence is closely related to working memory capacity. Here we aim to determine the associations of neural activation patterns during the *N*-back working memory paradigm with psychometric intelligence and working memory performance. We solved the statistical problems of previous studies using (1) a large cohort of 1235 young adults and (2) robust voxel-by-voxel permutation-based statistics at the whole-brain level. Many of the significant correlations were weak, and our findings were not consistent with those of previous studies. We observed that many of the significant correlations involved brain areas in the periphery or boundaries between the task-positive network (TPN) and task-negative network (TNN), suggesting that the expansion of the TPN or TNN is associated with greater cognitive ability. Lower activity in TPN and less task-induced deactivation (TID) in TNN were associated with greater cognitive ability. These findings indicate that subjects with greater cognitive ability have a lower brain response to task demand, consistent with the notion that TID in TNN reflects cognitive demand but partly inconsistent with the prevailing neural efficiency theory. One exception was the pre-supplementary motor area, which plays a key role in cognitive control and sequential processing. In this area, intelligent subjects demonstrated greater activity related to working memory, suggesting that the pre-supplementary motor area plays a unique role in the execution of working memory tasks in intelligent subjects.

Introduction

Psychometric intelligence is an important individual difference, as it can predict performance in a wide range of cognitive, academic, and job-related tasks (Cattell 1971). One way to measure psychometric intelligence is to use standard IQ tests that assess cognitive ability across multiple tasks and domains (Cattell 1971). Another method is to use tests of non-verbal reasoning (called tests of general fluid intelligence) in which one has to solve novel non-verbal problems (Cattell 1971). Indeed, these two measures strongly correlate (Cattell 1971). Working memory is the limited capacity storage system involved in the maintenance and manipulation of information over short periods of time. The working memory span correlates with performance on a wide range of tasks (Baddeley 2003) and is also strongly associated with general intelligence (Engle et al. 1999a). The neural mechanisms underlying individual differences in psychometric and general intelligence have been subject to intensive investigation (Jung and Haier 2007). The networks commonly activated during externally directed attention-demanding tasks are the lateral prefrontal cortex (LPFC) and posterior parietal cortex (e.g., the inferior/superior parietal lobule), together known as the task-positive network (TPN). This network is activated during working memory performance and non-verbal reasoning tasks (Baddeley 2003). Alternatively, the task-negative network (TNN) [or default mode network (DMN)], which includes the medial prefrontal cortex (mPFC), posterior cingulate gyrus (PCC), precuneus, and temporo-parietal junction, is commonly deactivated during externally directed attention demanding tasks such as working memory performance and non-verbal reasoning tasks (Baddeley 2003).

Many neuroimaging studies have investigated the neural mechanisms underlying individual differences in psychometric intelligence, including several

examining the association between psychometric intelligence and neural activation as measured by functional magnetic resonance imaging (fMRI) during the N-back working memory paradigm. However, several findings were inconsistent across these studies. Four such studies (Burgess et al. 2011; Gray et al. 2003; Tang et al. 2010; Waiter et al. 2009) reported differences in positive and negative correlations between psychometric intelligence and neural activation in the prefrontal, parietal, and other brain regions (Table 1). Another study (Basten et al. 2013) using multiple types of working memory tasks (including updating tasks, similar to the N-back task) found that more intelligent subjects exhibited greater combined brain activity in the fronto–parieto–temporal areas, hippocampus, and cerebellum during these tasks (Basten et al. 2013).

These discordant results could be due to any number of methodological differences (though this idea is speculation and not supported by existing theories). Moreover, these previous neuroimaging studies are limited by several unresolved statistical issues. First, several studies reporting significant correlations or that failed to find significant correlations did not correct for multiple comparisons in whole-brain analyses. For example, Gray et al. (2003) and Tang et al. (2010) used uncorrected P values. Alternatively, a study by Waiter et al. (2009) used the appropriate correction for multiple comparisons, specifically an SPM-based voxel-level correction for multiple comparison at the wholebrain level that was previously validated as an appropriate control for false positives (Eklund et al. 2016). This study found no significant association between general intelligence score and brain activity during the N-back working memory task. Basten et al. (2013) and Burgess et al. (2011) used the alphasim-based cluster size threshold in their studies, which was shown to be inappropriate for such applications (Eklund et al. 2016). In the latter study, procedures defining ROIs for

analyses of general intelligence included identification of voxels with significant correlations between brain activity and task accuracy, which in turn correlates with general intelligence. This kind of ‘double-dipping’ procedure has been shown to be improper (Vul et al. 2009).

The second problem with previous neuroimaging studies investigating the association between working memory capacity and general intelligence is small sample size ($N = 100$ at most). Recent studies have demonstrated several pitfalls of whole-brain analyses, especially whole-brain analyses of individual differences, thereby underscoring the importance of large cohorts. For example, recent studies have shown that the widely used cluster-based corrections for multiple comparisons at the whole-brain level included in a variety of software packages are liberal and lead to substantial numbers of false positives (Eklund et al. 2016). In addition, tests have shown that the reliability of inter-individual brain activity is not high, and even established simple-task paradigms such as the N-back task have a reliability of approximately 0.5 for specific ROIs (Plichta et al. 2012). The use of conservative corrections for multiple comparisons, such as voxel-level corrections for multiple comparisons in Statistical Parametric Mapping software (Eklund et al. 2016), requires a large sample size to overcome this low reliability. Without a large cohort, the results of these analyses are likely to be noncomprehensive, and the strength of the results will be highly overestimated (Vul et al. 2009) or more likely to yield false positives (Button et al. 2013). Some previous studies on general intelligence and working memory-related brain activity focused on ROIs (Tang et al. 2010). However, previous fMRI studies using a large cohort found that brain activity in areas that lie on the border between networks and some individuals show activation during the tasks while others show deactivation

during the tasks, are also robustly associated with performance or other psychometric measures (Mennes et al. 2010; Takeuchi et al. 2014a). In such cases, ROI analyses cannot properly identify the areas of significant associations between activity and cognitive performance.

The third problem is the unresolved issue of whether the positive and negative correlations of performance with brain activity represent associations with activation or with deactivation. In other words, it is unclear whether a positive correlation between cognitive performance and brain activity represent the association between better cognitive performance and greater neural activation or less deactivation. Basten et al. (2013) observed that general intelligence positively correlated with brain activity on average, but suggested that more intelligent subjects exhibited greater activity in TPN with less deactivation in TNN. No other studies have addressed this issue and this question remains unresolved due to the aforementioned statistical problems.

Despite the importance of general intelligence and working memory, there is still no comprehensive picture of the associations between brain activity during working memory and general intelligence. The purpose of this study is to address the three aforementioned problems and to identify brain regions in which working memory-related brain activity is significantly associated with general intelligence and working memory span. This question is investigated at the whole-brain level using corrections for multiple comparisons shown to properly control for false positives.

Through these investigations, we tested three major important hypotheses or questions in this field. The first question is whether performance and activity correlate in (a) TPN, (b) TNN, and (or) (c) border areas, as described above (here TPN and TNN are defined as the areas that are activated and deactivated during the externally directed

attention-demanding tasks used in the present study). The second hypothesis to be tested relates to neural efficiency theory, which is based on diverse neuroimaging results. The neural efficiency theory is based on observations that intelligent people generally show lower prefrontal activity during tasks of low-to-moderate difficulty or after sufficient practice to develop efficient strategies compared to less intelligent people (Neubauer and Fink 2009), while in more complex tasks, there is a positive correlation between brain activity and general intelligence (Neubauer and Fink 2009). A review of this topic highlighted the importance of parietal lobes in intelligence (Jung and Haier 2007). Similarly, a review by Toffanin et al. (2007) suggested that highly intelligent subjects are more likely to recruit the parietal lobes. To study this hypothesis, we set relatively easy task conditions and allowed sufficient practice with instructions regarding strategies, and then tested whether intelligence negatively correlates with activity in the prefrontal areas and positively with parietal lobe activity.

The third question addressed asks whether subjects with greater cognitive ability show more or less task-induced deactivation (TID) in TNN. According to the view that TID reflects more efficient attentional reallocation (Sambataro et al. 2008), subjects with greater fluid intelligence should show more TID. Consistent with this view is the finding that an elderly cohort exhibited both less TID in TNN and a substantial decline in fluid intelligence (Park et al. 2002). On the other hand, according to the view that TID reflects cognitive load (McKiernan et al. 2003), subjects with greater cognitive ability should show less TID. While the findings of Basten et al. (2013) are consistent with this view, such findings have not been reported in other studies.

Material and Methods

Subjects

The present study, which is a part of an ongoing project to investigate the associations among brain imaging, cognitive function, and aging, included relevant reliable cognitive measures and imaging data from 1235 healthy righthanded individuals (710 males and 525 females). The mean age of the subjects was 20.8 years [standard deviation (SD) 1.8 years, age range 18–27 years]. For detailed subject information, see Supplemental Methods. See Supplemental Discussion for the limitations conferred by this choice of cohort. Written informed consent was obtained from all participants or their guardians. This study was approved by the Ethics Committee of Tohoku University.

Assessment of psychometric measures of general intelligence

Raven's Advanced Progressive Matrix (Raven 1998), a measure of non-verbal reasoning shown to be most strongly correlated with general intelligence (Raven 1998), was used to assess psychometric intelligence. This test contains 36 non-verbal items requiring fluid reasoning ability. Each item consists of a 3×3 matrix with a missing piece to be completed by selecting the best among eight alternatives. The score on this test (number of correct answers in 30 min) was used as an index of individual psychometric intelligence. The description of this test is reproduced from our previous study that used the same assessment methods (Takeuchi et al. 2015b).

Assessment of working memory span

Computerized forward and backward digit span tests were used to assess working memory span. Subjects were asked to view a progressively increasing number

of random digits visually presented at 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 incorrectly responded 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. The following descriptions are reproduced from our previous study using the same method (Takeuchi et al. 2011a).

Some studies have emphasized the difference between simple working memory span tasks, such as the letter span task, and complex working memory tasks, such as the reading span task, which is based on a greater association of the latter span tasks with other cognitive functions. In this study, we selected a simple span task (digit span tasks) instead of complex span tasks such as reading span. This choice was made based on the view that the difference in complexity is only important for the cognitive functions relevant to the operation task used in the complex working memory task (such as reading in reading span tasks) (Wechsler 1997) (although, this is a matter of controversy).

Also, although some studies have stressed the difference between the span-forward task and the span-backward task, others deny the importance of this difference (Wechsler 1997). We used a standard combined score of the span-forward and span-backward tasks based on the latter view.

fMRI task.

fMRI was used to map brain activity during cognitive tasks. The descriptions

of this task are reproduced from our previous study using the same methods (Takeuchi et al. 2015a). The N-back task is a typical task for fMRI studies. We used the N-back task with conditions of 0-back (simple cognitive processes) and 2-back (working memory). We used a simple block design and the N-back working memory task (Callicott et al. 1999) to map brain activity during working memory. The N-back task was performed during fMRI scanning as described in our previous studies (Takeuchi et al. 2011a, c, 2014b). For more details, see Supplemental Methods. Sufficient practice was allowed, and we ascertained that subjects understood the tasks and the strategy of updating items to remember two by two during the 2-back task (Takeuchi et al. 2012). Several previous studies investigated brain activity related to the trial of lure during the N-back task (Burgess et al. 2011; Gray et al. 2003). A lure trial is the trial in which the stimuli which are correct answers in contingent trials appear and are therefore confusing. Since only four stimuli were used in this study, the whole block was regarded as related to the lures. Reaction time (RT) and accuracy on 0-back and 2-back tasks were used in analyses.

We have previously conducted experiments using N-back tasks on a similar subject sample (e.g., Takeuchi et al. 2012) and the measurement parameters used here, including task difficulty, were chosen based on the results and experiences in such experiments. We believe that the task was not too easy for testing our hypotheses despite the almost 100% accuracy (see “Results”). First, in previous neuroimaging studies of brain activities during N-back tasks in psychiatric patients and controls, the ease of the tasks did not prevent clear group differences in brain activities. For example, Jansma et al. (2004) demonstrated brain activity differences between patients and controls during the 0-back task despite almost 100% accuracy in both groups. Further,

group differences were similar to those of the 1-back task, in which patients made some mistakes. The same pattern was observed in another study comparing brain activities during the 0-back task and 2-back task between controls and relatives of schizophrenics (Whitfield-Gabrieli et al. 2009). But in Jansma et al. (2004), during the more difficult 2-back and 3-back tasks, such differences of activity became smaller in ROIs due to activity reduction in patients. Therefore, the assumption that more difficult N-back tasks can better reveal activity differences is not congruent with previous findings. Second, as described in the “Introduction”, neural efficiency theory itself hypothesizes differences of brain activity even when the task is easy (Neubauer and Fink 2009), and our study is testing predictions of the neural efficiency theory, so the tasks are supposed to be easy.

Image acquisition.

MRI data acquisition was conducted using a 3T Philips Achieva scanner. Forty-two transaxial gradient-echo images (echo time, 30 ms; flip angle, 90°; slice thickness, 3 mm; FOV, 192 mm; matrix, 64 × 64) covering the entire brain were acquired at a repetition time of 2.5 s using an echo-planar sequence. For the N-back session, 174 functional volumes were obtained. Diffusion-weighted data were acquired using a spin-echo echo-planar imaging (EPI) sequence (TR, 10293 ms; TE, 55 ms; FOV, 22.4 cm; 2 × 2 × 2 mm³ voxels; 60 slices; SENSE reduction factor, 2; number of acquisitions, 1). The diffusion weighting was isotropically distributed along 32 directions (*b* value, 1000 s/mm²). Additionally, using a spin-echo EPI sequence (TR, 10293 ms; TE, 55 ms; FOV, 22.4 cm; 2 × 2 × 2 mm³ voxels; 60 slices), images with no diffusion weighting (*b* value, 0 s/mm²) (*b* = 0 images) were acquired. From the collected images, fractional anisotropy (FA) and mean diffusivity (MD) maps were calculated (Takeuchi et al.,

2011b) and used for preprocessing of BOLD images (see Supplemental Methods for details). The descriptions of this subsection are mostly reproduced from our previous study using the exact same methods (Takeuchi et al., 2015a).

Preprocessing of imaging data

Preprocessing and analysis of functional activation data were performed using SPM8 implemented in Matlab. Here we provide a summary, while the full details and methodological considerations are provided in the **Supplemental Methods**. Before analysis, individual BOLD images were realigned and re-sliced to the mean BOLD image, which was then realigned to the mean $b = 0$ image as previously described (Takeuchi et al., 2011a). 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. All images were normalized using a previously validated two-step “new segmentation” algorithm of diffusion images and the previously validated diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL)-based registration (Takeuchi et al., 2013a). The voxel size of normalized BOLD images was $3 \times 3 \times 3$ mm³. The descriptions in this subsection are mostly reproduced from our previous study using the exact same methods (Takeuchi et al., 2015a).

First-level analysis of functional imaging data

Individual-level statistical analyzes were performed using a general linear model. A design matrix was fitted to each participant with one regressor in each N-back task condition (0- or 2-back) using a standard hemodynamic response function. The cue phases of the N-back task were modeled in the same manner but were not analyzed

further. Six parameters obtained by rigid body correction of head motion were regressed out by including these parameters in the regression model. The design matrix weighted each raw image according to its overall variability to reduce the impact of movement artifacts (Diedrichsen and Shadmehr, 2005). We removed low-frequency fluctuations using a high-pass filter with a cutoff value of 128 seconds. After estimation, beta images were smoothed (8 mm full-width at half-maximum) and taken to the second level of analysis. The descriptions in this subsection are mostly reproduced from our previous study using the exact same methods (Takeuchi et al., 2015a).

Statistical analyzes of behavioral data

The behavioral data were analyzed using SPSS 22.0 statistical software (SPSS Inc., Chicago, IL). In each analysis, $P < 0.05$ was considered statistically significant. Associations among RAPM score, digit span score, and RTs on the 0-back and 2-back tasks were analyzed using simple regression.

Group-level whole-brain imaging data analysis.

At the group level, we tested for relationships between individual cognitive ability and regional brain activity levels during the 0-back and 2-back tasks as well as relationships between individual cognitive ability and working memory-specific regional activity (2-back –0-back contrast). In the whole-brain analysis, we used multiple linear regression to identify areas where the beta estimates of the contrast were significantly related to individual cognitive ability. The effects of sex, age, accuracy, and RTs of the 2-back task and 0-back task, and volume-level mean framewise displacement during the scan for the N-back task (Power et al., 2012) were corrected for by entering these parameters into

the multiple regression model as covariates. Each whole-brain analysis included only one cognitive ability of interest. Therefore, we performed six brain analyzes (3 contrasts [0-back task, 2-back task, and 2-back–0-back] \times 2 tasks [RAPM and digit span] = 6).

A multiple comparison correction was performed using the T score with randomized (5,000 permutations) nonparametric testing using the toolbox (<http://dbm.neuro.uni-jena.de/tfce/>). We applied a voxel threshold of FWE corrected at $P < 0.05$.

Accuracy and RT were adjusted in the analyzes. While this is considered the gold standard for such analyzes, we are aware that some of these adjustments may regress out some relevant components of general intelligence and working memory capacity (such as cognitive speed). However, without this adjustment, it is unclear whether the observed differences in neural activity simply reflect behavioral differences themselves.

Results

Basic psychological data

Mean (\pm SD) age, RAPM score, digit span score, accuracies and RTs of the 0-back and 2-back tasks, and volume-wise framewise displacement are presented in Supplemental Table 1.

Response accuracies on the 2-back and 0-back tasks showed ceiling effects ($>99.0\%$ correct on average). Significant associations were observed among greater RAPM score, greater digit span score, lower 0-back task RT, and lower 2-back task RT (Table 2). The correlation between the 0-back task RT and 2-back task RT was strong ($r = 0.612$), while the other correlations were weak ($r < 0.3$). Ceiling effects are not likely

to diminish the associations between brain activity and cognitive abilities because even brain activities during easy tasks can show associations with cognitive abilities, a finding observed in this study as well.

Nonetheless, the weak association between working memory capacity and general intelligence (RAPM score) might still reflect the choice of tasks. However, in college cohorts, this level of correlation between working memory performance and general intelligence as measured by non-verbal reasoning (RAPM score) is commonly observed even when a complex working memory span task is used to measure working memory capacity (Engle et al., 1999b).

Correlation between brain activity during the N-back task and performance on a general intelligence task

After controlling for confounding variables, multiple regression analysis revealed that the RAPM score was significantly and positively correlated with brain activity level during the 0-back task (0-back > rest) in parts of the superior frontal gyrus defined by automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002) [the right rostralateral prefrontal cortex (RLPFC)], which is deactivated during the 0-back task (Fig 1a, Table 3). There was a significant negative correlation of RAPM score with brain activity during the 0-back task in areas spread across the right inferior parietal lobule, right supramarginal gyrus, right postcentral gyrus, and lateral areas of the superior frontal gyrus in AAL (right premotor cortex), which are activated during the 0-back task (Fig 1b,c, Table 3).

After controlling for confounding variables, multiple regression analysis revealed a significant positive correlation between RAPM score and brain activity

during the 2-back task (2-back > rest) in areas spread across the precuneus and cuneus (AAL labels), areas deactivated during the 2-back task (Fig. 2a, Table 3). There was also a significant positive correlation in part of the supplementary motor area (AAL label), specifically in the pre-supplementary motor area (the region rostral to the vertical plane through the anterior commissure) ((Picard and Strick, 1996; Zilles et al., 1995), which was activated during the 2-back task (Fig. 2b, Table 3). Finally, there was a significant positive correlation in an area spread across the superior frontal medial area and lateral area of the superior frontal gyrus (AAL labels) (the dorsal part of the mPFC), which was deactivated during the 2-back task (Fig. 2c, Table 3). A significant negative correlation was observed between the RAPM score and brain activity during the 2-back task in areas spread across the right inferior parietal lobule, right supramarginal gyrus, and right postcentral gyrus (AAL labels), which were activated during the 2-back task (Fig. 2d) (Table 3).

After controlling for confounding variables, multiple regression analysis revealed a significant positive correlation between RAPM score and brain activity specific to working memory (2-back–0-back) in an area mainly within the right hippocampus, but also covering the right parahippocampal gyrus and right amygdala (AAL labels), which were deactivated (2-back–0-back), and the supplementary motor area (AAL label) or the pre-supplementary motor area, which was activated (2-back–0-back) (Fig. 3a,b) (Table 3).

Correlation between brain activity during the N-back task and a working memory span task

After controlling for confounding variables, multiple regression analysis

revealed a significant positive correlation between digit span score and brain activity during the 0-back task in an area within the medial part of the superior frontal gyrus (AAL label) (the dorsomedial prefrontal cortex), which was neither activated nor deactivated during the 0-back task (Fig 4) (Table 4).

There were no significant associations between digit span score and brain activity during the 2-back task.

After controlling for confounding variables, multiple regression analysis revealed a significant positive correlation between digit span score and brain activity specific to working memory (2-back–0-back) in an area spread across the right hippocampus and the right parahippocampal gyrus (AAL labels), which was deactivated in this contrast (Fig. 5a) (Table 4), areas of the left postcentral gyrus, which were deactivated in this contrast (Fig. 5b) (Table 4), the right Rolandic operculum, which was deactivated in this contrast (Fig. 5c) (Table 4), and the left parahippocampal gyrus, which was deactivated (Fig. 5a) (Table 4).

Discussion

The present study newly reveals associations between regional brain activity during working memory and general intelligence in a very large cohort ($N > 1200$) using robust permutation-based voxel-by-voxel statistics at the whole-brain level to prevent the problems described in the “Introduction”. We tested neural efficiency theory and two theories involving individual differences of TID in TNN, and investigated the neural locations of significant correlations (positive or negative correlations, TPN, TNN, border areas).

The simple correlation coefficients between beta estimates of brain activity in

areas of significance and RAPM score or digit span were relatively small (around 0.1–0.16) in the present study. Further, when voxel-level correction for multiple comparisons by permutation was used, the areas of significance were limited despite the huge sample size. These small correlations may reflect a relative weak association between task-dependent brain activation pattern and cognitive ability, as more widespread associations were observed when an uncorrected threshold $P < 0.001$ was applied. Therefore, this pattern of findings may have same roots (the effect is small). Relatively weak correlations ($r < 0.2$) between individual cognitive differences and neuroimaging measures are actually common in studies using huge samples of young normal adults (i.e., $N > \text{several hundred}$) regardless of the specific imaging measures (Magistro et al. 2015; Schilling et al. 2012; Takeuchi et al. 2015c, 2017). In fact, we are not aware of any exceptions. The large effect sizes and correlation coefficients for peak voxels in significant areas yielded by whole-brain analyses on small samples do not indicate true effect sizes or correlation strengths (Murphy et al. 2012). Rather, in whole-brain analyses, especially those on small samples, the effect sizes were overestimated due to overfitting (Vul et al. 2009). In addition, due to publication bias, and stringent thresholds, the studies of whole-brain analyses with small sample size cannot report the significant findings of small effect size. Even when single studies show remarkable effect size (especially under low statistical power), the effect size reported in meta-analyses can be very small (e.g., Murphy et al. 2012). Adding these up, we cannot presume the large effect size in these kinds of studies.

From the observed weak correlations, we believe that it is critical to use relatively large sample sizes for this kind of study and that such large samples increase the reliability of the results. To detect a small effect such as $r = 0.1$ reliably (e.g., $\alpha =$

0.001; power = 50%), a sample size of approximately 1080 subjects is required.

Therefore, the present results validate the necessity for a large sample size. Statistical power is typically very low in the field of neuroscience (Button et al. 2013), which leads not only to overestimates of effect size but also low reproducibility. Indeed, the general replicability of combined neuroimaging/psychometric studies was recently questioned (Eklund et al. 2016). Therefore, in addition to novelty, increased reliability, which requires larger sample size, is critical.

Several factors may account for the weak correlation coefficients reported in the present study. It is important to note, however, that weak correlation coefficients do not necessarily indicate that the associations between brain activities that fMRI is expected to measure and cognitive functions are weak. Rather, factors such as low reliability of fMRI measures during N-back tasks (around 0.5 according to Plichta et al. 2012), differences in subject condition, image preprocessing inaccuracies, and random factors such as temperature, noise in the scanner, or the subject's specific thoughts during rest may undermine the strengths of the associations between cognitive function and fMRI measures. Regardless, the weak correlations suggest limitations of fMRI for purposes such as predicting an individual's cognitive capacity.

Many of the significant associations we observed between cognitive function and brain activities were at the border between areas activated or deactivated during the task or at the periphery of such areas (i.e., areas distant from those where most subjects showed peak activation or deactivation during the task). The findings indicate in many of the cases, subjects with greater cognitive abilities have more/less additive peripheral areas of the functional areas or networks in the border areas of different functional network areas, rather than greater/lesser strength of the functional activation or

deactivation among the centers of the nodes in each functional network. This pattern is congruent with our previous findings of associations between resting-state functional connectivity and cognitive differences in large samples (Takeuchi et al. 2013b, 2014a) and may thus be a robust feature of associations between cognitive differences and functional imaging indices. A previous study reported similar brain map pattern, with concentrations of correlations in border areas between TPN and TNN (Basten et al. 2013). Whether these areas are activated or deactivated during certain tasks is likely to be unaffected by variables such as mood, thoughts, and actions during the task and noise during the scan. Thus, these measures are likely stable and sensitive. The concentrations of the associations in these kinds of peripheral regions may be another reason why significant associations were limited to small areas.

The predictions based on neural efficiency theory (Neubauer and Fink 2009) that people with high intelligence will exhibit lower prefrontal activity during tasks of low to moderate difficulty and stronger dependence on the parietal lobe were not confirmed by the present study. Both the 0-back and 2-back tasks were easy enough as evidenced by high (near perfect) response accuracy, primarily due to careful explanation of the task and practice time allowed prior to scanning. The correlations between cognitive function and brain activity under higher load (0-back > rest, 2-back > rest, 2-back–0-back) were mostly positive in the prefrontal cortex (Figs. 1a, 2b, 3b, 4) except for right premotor cortex (Fig. 1c). Further, the correlations between cognitive function and brain activity under higher load (0-back > rest, 2-back > rest, 2-back–0-back) were mostly negative in the parietal lobe (Figs. 1b, 2d) except for the precuneus (Fig. 2a). Therefore, incongruent with predictions of the neural efficiency hypothesis, both positive and negative correlations were observed in the frontal and parietal cortices. As

this theory was devised using inductive reasoning, the principles of the theory may require revision as more experimental evidence obtained by robust statistics is being accumulated.

The present results are congruent with the hypothesis that deactivation in the TNN reflects the cognitive load. Positive associations between cognitive function and brain activity associated with higher load (0-back > rest, 2-back > rest, 2-back-0-back) were observed in the precuneus and dorsal part of the mPFC, key nodes of the TNN (DMN) (Buckner et al., 2008), as well as in areas deactivated during the tasks such as the hippocampus, contingent regions and right RLPFC (Fig. 1). The degree of deactivation in the TNN (DMN) is thought to reflect the cognitive load (McKiernan et al., 2003) or the degree of concentration applied during the task (Sambataro et al., 2008). The present results are consistent with a previous study showing that subjects with higher IQ exhibit less deactivation during cognitive tasks (Basten et al., 2013). Thus, the present results suggest that subjects with greater cognitive ability face a lower cognitive load during cognitive tasks and exhibit relatively higher brain activity (less deactivation) in these regions.

With the exception of the pre-supplementary motor area (Figs. 2b, 3b) and border areas between the networks (Fig. 4), subjects with greater general intelligence or working memory span showed lower levels of activation and deactivation (Figs. 1, 2, 3, 5). In other words, subjects with greater cognitive ability exhibited lower activity increases or less deactivation as task demand increased. Therefore, the present results are not in perfect accord with the neural efficiency theory, which suggests that subjects with greater intelligence should exhibit lower brain activation in the prefrontal area (and high activation in the parietal area). However, the notion that subjects with greater

intelligence execute tasks more efficiently and face less cognitive demand appears consistent with our findings. Alternatively, subjects with greater intelligence may maintain greater brain activity during rest. Additional studies are needed to determine whether this pattern appears when using other functional imaging paradigms.

We observed that working memory-related brain activity (during the 2-back task) and working memory-specific brain activity (2-back–0-back contrast) in the pre-supplementary motor area correlated significantly and positively with general intelligence. This area is continuously activated during working memory tasks, with greater activation as the memory load increases (Owen et al. 2005). The pre-supplementary motor area is preferentially involved in higher-level planning processes and more abstract aspects of cognition, while the supplementary motor area proper (the more caudal region) is strictly associated with motor execution (Cona and Semenza 2017; Picard and Strick 1996; Seitz et al. 2006). The pre-supplementary motor area contributes to sequence operations in a variety of cognitive domains (Cona and Semenza 2017), aspects of task switching, and changing of plans in cognitive control (Nachev et al. 2008). All of these processes are likely to be involved in the effective execution of N-back tasks, which require remembering sequential information and continuous updating. Intelligent subjects may recruit this area and activate these cognitive processes, thereby reducing cognitive load.

The present study newly reveals the associations among regional brain activity during working memory tasks, general intelligence, and working memory span using rigorous statistics and a large cohort, thereby solving statistical problems that have limited the reliability of previous studies. Many findings of the current study were not observed in previous studies. Many of the significant associations between cognitive

functions and brain activity were found in the border regions between areas activated or deactivated during the tasks or at the periphery of such areas, suggesting that the expansion of activated or deactivated areas is associated with higher cognitive function. The strengths of all correlations were weak. Further, many of our findings were not predicted by the neural efficiency theory well. We observed that less activity in TPN and less TID in TNN were associated with greater cognitive ability, indicating that subjects with greater cognitive ability show a smaller change in brain activity or deactivation under higher task demand. This finding is consistent with the idea that TID reflects the cognitive demand. One exception was the pre-supplementary motor area, which plays a key role in cognitive control and sequential processing. In this area, intelligent subjects showed greater activity related to working memory, suggesting that this area plays a unique role in the execution of working memory performance in intelligent subjects.

Acknowledgments

We thank Yuki Yamada for operating the MRI scanner, Haruka Nouchi for conducting the psychological tests, all other assistants for helping with the experiments and the study, the study participants, and all our other colleagues at IDAC, Tohoku University, for their support. This study was supported by JST/RISTEX, JST/CREST, a Grant-in-Aid for Young Scientists (B) (KAKENHI 23700306) and a Grant-in-Aid for Young Scientists (A) (KAKENHI 25700012) from the Ministry of Education, Culture, Sports, Science, and Technology. The authors would like to thank Enago (www.enago.jp) for the English language review.

Ethical statement

Compliance with Ethical Standards: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Funding: This study was supported by JST/RISTEX, JST/CREST, a Grant-in-Aid for Young Scientists (B) (KAKENHI 23700306) and a Grant-in-Aid for Young Scientists (A) (KAKENHI 25700012) from the Ministry of Education, Culture, Sports, Science, and Technology.

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: This study was approved by the Ethics Committee of Tohoku University.

Informed consent: Informed consent was obtained from all individual participants included in the study.

References

- Baddeley A (2003) Working memory: looking back and looking forward. *Nat Rev Neurosci* 4:829-839.
- Basten U, Stelzel C, Fiebach CJ (2013) Intelligence is differentially related to neural effort in the task-positive and the task-negative brain network. *Intelligence* 41:517-528.
- Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network. *Ann N Y Acad Sci* 1124:1-38.
- Burgess GC, Gray JR, Conway AR, Braver TS (2011) Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *J Exp Psychol Gen* 140:674-692.
- Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14:365-376.
- Callicott JH, Mattay VS, Bertolino A, Finn K, Coppola R, Frank JA, Goldberg TE, Weinberger DR (1999) Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex* 9:20-26.
- Cattell RB (1971) *Abilities: Their structure, growth, and action*. Boston: Houghton-Mifflin.
- Cona G, Semenza C (2017) Supplementary motor area as key structure for domain-general sequence processing: A unified account. *Neurosci Biobehav Rev* 72:28-42.
- Diedrichsen J, Shadmehr R (2005) Detecting and adjusting for artifacts in fMRI time series data. *Neuroimage* 27:624-634.

- Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*:201602413.
- Engle RW, Kane MJ, Tuholski SW (1999a) Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In: *Models of working memory: Mechanisms of active maintenance and executive control* (Miyake A, Shah P, eds), pp 102-134. Cambridge: Cambridge University Press.
- Engle RW, Tuholski SW, Laughlin JE, Conway AR (1999b) Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J Exp Psychol Gen* 128:309-331.
- Gray JR, Chabris CF, Braver TS (2003) Neural mechanisms of general fluid intelligence. *Nat Neurosci* 6:316-322.
- Jansma J, Ramsey N, Van Der Wee N, Kahn R (2004) Working memory capacity in schizophrenia: a parametric fMRI study. *Schizophr Res* 68:159-171.
- Jung RE, Haier RJ (2007) The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav Brain Sci* 30:135-154.
- Magistro D, Takeuchi H, Nejad KK, Taki Y, Sekiguchi A, Nouchi R, Kotozaki Y, Nakagawa S, Miyauchi CM, 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 JA, Laurienti PJ, Burdette JH (2004) Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage* 21:450-455.

- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19:1233-1239.
- McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR (2003) A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J Cogn Neurosci* 15:394-408.
- Mennes M, Kelly C, Zuo XN, Di Martino A, Biswal BB, Castellanos FX, Milham MP (2010) Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. *Neuroimage* 50:1690-1701.
- 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. *Mol Psychiatry* 18:512-520.
- Nachev P, Kennard C, Husain M (2008) Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience* 9:856-869.
- Neubauer AC, Fink A (2009) Intelligence and neural efficiency. *Neurosci Biobehav Rev* 33:1004-1023.
- Owen AM, McMillan KM, Laird AR, Bullmore E (2005) N - back working memory paradigm: A meta - analysis of normative functional neuroimaging studies. *Hum Brain Mapp* 25:46-59.
- Park DC, Lautenschlager G, Hedden T, Davidson NS, Smith AD, Smith PK (2002) Models of visuospatial and verbal memory across the adult life span. *Psychol Aging* 17:299-320.
- Picard N, Strick PL (1996) Motor areas of the medial wall: a review of their location and functional activation. *Cereb Cortex* 6:342-353.

- Plichta MM, Schwarz AJ, Grimm O, Morgen K, Mier D, Haddad L, Gerdes A, Sauer C, Tost H, Esslinger C (2012) Test–retest reliability of evoked BOLD signals from a cognitive–emotive fMRI test battery. *Neuroimage* 60:1746-1758.
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59:2142-2154.
- Raven J (1998) Manual for Raven's progressive matrices and vocabulary scales. Oxford: Oxford Psychologists Press.
- Sambataro F, Murty VP, Callicott JH, Tan HY, Das S, Weinberger DR, Mattay VS (2008) Age-related alterations in default mode network: Impact on working memory performance. *Neurobiol Aging*.
- 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. *Mol Psychiatry* 18:624-630.
- Seitz RJ, Nickel J, Azari NP (2006) Functional modularity of the medial prefrontal cortex: involvement in human empathy. *Neuropsychology* 20:743.
- Takeuchi H, Taki Y, Hashizume H, Sassa Y, Nagase T, Nouchi R, Kawashima R (2011a) Failing to deactivate: the association between brain activity during a working memory task and creativity. *Neuroimage* 55:681-687.
- Takeuchi H, Taki Y, Sassa Y, Hashizume H, Sekiguchi A, Fukushima A, Kawashima R (2011b) Verbal working memory performance correlates with regional white matter structures in the fronto-parietal regions. *Neuropsychologia* 49:3466-3473
- Takeuchi H, Taki Y, Hashizume H, Sassa Y, Nagase T, Nouchi R, Kawashima R (2011c)

- Effects of training of processing speed on neural systems. *J Neurosci* 31:12139-12148.
- Takeuchi H, Sugiura M, Sassa Y, Sekiguchi A, Yomogida Y, Taki Y, Kawashima R (2012) Neural correlates of the difference between working memory speed and simple sensorimotor speed: an fMRI study. *PLoS ONE* 7:e30579.
- Takeuchi H, Taki Y, Thyreau B, Sassa Y, Hashizume H, Sekiguchi A, Nagase T, Nouchi R, Fukushima A, Kawashima R (2013a) White matter structures associated with empathizing and systemizing in young adults. *Neuroimage* 77:222-236.
- Takeuchi H, Taki Y, Nouchi R, Sekiguchi A, Hashizume H, Sassa Y, Kotozaki Y, Miyauchi CM, Yokoyama R, Iizuka K (2013b) Resting state functional connectivity associated with trait emotional intelligence. *Neuroimage* 83:318-328.
- Takeuchi H, Taki Y, Nouchi R, Sekiguchi A, Hashizume H, Sassa Y, Kotozaki Y, Miyauchi CM, Yokoyama R, Iizuka K (2014a) Association between resting-state functional connectivity and empathizing/systemizing. *Neuroimage* 99:312-322.
- Takeuchi H, Tomita H, Taki Y, Kikuchi Y, Ono C, Yu Z, Sekiguchi A, Nouchi R, Kotozaki Y, Nakagawa S (2015a) 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.
- Takeuchi H, Taki Y, Sekiguchi A, Nouchi R, Kotozaki Y, Nakagawa S, Miyauchi CM, Iizuka K, Yokoyama R, Shinada T (2015b) Amygdala and cingulate structure is associated with stereotype on sex-role. *Scientific reports* 5:article 14220.
- Takeuchi H, Taki Y, Nouchi R, Hashizume H, Sassa Y, Sekiguchi A, Kotozaki Y, Nakagawa S, Nagase T, Miyauchi M, C., Kawashima R (2014b) Associations

- among imaging measures (2): The association between gray matter concentration and task-induced activation changes. *Hum Brain Mapp* 35:185-198.
- Takeuchi H, Taki Y, Nouchi R, Sekiguchi A, Hashizume H, Sassa Y, Kotozaki Y, Miyauchi CM, Yokoyama R, Iizuka K, Seishu N, Tomomi N, Kunitoki K, Kawashima R (2015c) 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, 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.
- Tang C, Eaves E, Ng J, Carpenter D, Mai X, Schroeder D, Condon C, Colom R, Haier R (2010) Brain networks for working memory and factors of intelligence assessed in males and females with fMRI and DTI. *Intelligence* 38:293-303.
- Toffanin P, Johnson A, De Jong R, Martens S (2007) Rethinking neural efficiency: effects of controlling for strategy use. *Behav Neurosci* 121:854-870.
- 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”.

Perspect Psycholo Sci 4:319-324.

Waiter GD, Deary IJ, Staff RT, Murray AD, Fox HC, Starr JM, Whalley LJ (2009) Exploring possible neural mechanisms of intelligence differences using processing speed and working memory tasks: An fMRI study. *Intelligence* 37:199-206.

Wechsler D (1997) *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Whitfield-Gabrieli S, Thermenos HW, Milanovic S, Tsuang MT, Faraone SV, McCarley RW, Shenton ME, Green AI, Nieto-Castanon A, LaViolette P (2009) Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *Proc Natl Acad Sci U S A* 106:1279-1284.

Zilles K, Schlaug G, Matelli M, Luppino G, Schleicher A, Qü M, Dabringhaus A, Seitz R, Roland P (1995) Mapping of human and macaque sensorimotor areas by integrating architectonic, transmitter receptor, MRI and PET data. *J Anat* 187:515.

Table 1. Summary of documented associations between general intelligence and brain activity during N-back tasks

Study	Psychological measure	Subjects	Activity measure	Brain regions
Burgess et al. (2011)	A factor from RAPM and CCFT	102 healthy young adults (58 females; mean age, 22.21 years; SD, 4.78 years)	Lure-related activities in the 3-back task of faces and words	ROI ^x : Positive correlations in bilateral DLPFC, mPFC, bilateral median and lateral parietal areas, left fusiform gyrus, and bilateral parahippocampal gyrus
Gray et al. (2003)	RAPM score	60 healthy adults (31 females; age, 18–37 years)	Lure-related activities in the 3-back task of faces and words	Uncorrected whole-brain analyzes: bilateral positive correlations in LPFC, dorsal ACC, right cerebellum, supramarginal gyrus, bilateral superior temporal gyrus
Tang et al. (2010)	General intelligence calculated from average of z scores of eight tests	40 healthy adults (19 females; mean age, 26.6 years; SD, 4.60 in males, 4.90 in females)	1,2,3-back vs. 0-back in the letter N-back task	ROI: negative correlations in the right parietal cortex
Waiter et al. (2009)	RPM ^y	52 of 79 older adults (28 females; age, 69–70 years)	2-back–0-back in the matching type letter N-back task	No significant results

CCFT, Cattell Culture Faire test; ROI, Region of interest (for analysis); RAPM, Raven’s Advanced Progressive Matrices; RPM, Raven’s

Progressive Matrices

^xNote: ROIs in this study included identification of voxels with a significant correlation between brain activity and task accuracy. This procedure has been shown to be improper for ROI selection to identify neural correlates of psychological measures that correlate with task accuracy (Vul et al., 2009).

Table 2. Associations among RAPM score, digit span score, and reaction times (RTs) of the 0-back and 2-back tasks (simple correlation coefficient, P-value)

	RAPM score	Digit span score	0-back RT	2-back RT
RAPM score	-	-	-	-
Digit span score	0.279, 2.05×10^{-23}	-	-	-
0-back RT	-0.120, 2.50×10^{-5}	-0.179, 2.34×10^{-10}	-	-
2-back RT	-0.194, 6.57×10^{-12}	-0.178, 3.22×10^{-10}	0.612, 1.02×10^{-127}	-

Table 3. Brain regions with significant correlations between brain activity and RAPM score

Included gray matter areas* (number of significant voxels in the left and right sides of each anatomical area)	x	y	z	T score	Corrected <i>P</i> -value (FWE)	Cluster size (voxels)	<i>r</i> **
<i>Positive correlation between 0-back activity and RAPM score</i>							
Superior frontal gyrus (other parts***) (R:4)	24	69	9	4.37	0.028	5	0.123
<i>Negative correlation between 0-back activity and RAPM score</i>							
Inferior parietal lobule (R:4)/Postcentral gyrus (R:13)/Supramarginal gyrus (R:9)	36	-36	45	4.84	0.006	26	-0.134
Superior frontal gyrus (other parts****) (R:1)	27	-12	57	4.36	0.035	1	-0.110
<i>Positive correlation between 2-back activity and RAPM score</i>							
Cuneus (R:10)/Precuneus (L:1, R:16)	6	-72	30	5.05	0.003	27	0.140

Pre-supplemental motor area (R:12)	9	9	69	4.74	0.008	12	0.123
Superior frontal gyrus (medial part) (R:1)/Superior frontal gyrus (other parts****) (R:1)							0.122
	12	42	54	4.38	0.036	3	
<i>Negative correlation between 2-back activity and RAPM score</i>							
Inferior parietal lobule (R:7)/Postcentral gyrus (R:14)/Supramarginal gyrus (R:5)							-0.150
	36	-39	48	4.83	0.008	26	
<i>Positive correlation between (2-back-0-back) activity and RAPM score</i>							
Amygdala (R:1)/Hippocampus (R:11)/Parahippocampal gyrus (R:3)							0.159
	18	-6	12	4.56	0.018	16	
Pre-supplemental motor area (R:1)	6	6	69	4.29	0.043	1	0.099

● *Labelings of the anatomical regions of gray matter were based on the WFU PickAtlas Tool (<http://www.fmri.wfubmc.edu/cms/software#PickAtlas/>) (Maldjian et al., 2004; Maldjian et al., 2003) and on 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.

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

*** This area is classified as the supplementary motor area in this atlas. However, as noted in the Results section, it has been pointed out this area specifically corresponds to pre-supplementary motor area ((Picard and Strick, 1996; Zilles et al., 1995), therefore, we called so in this table, too.

**** Areas of the superior frontal gyrus other than the medial, orbital, and medial-orbital parts.

Table 4. Brain regions with significant correlations between brain activity and digit span score

Included gray matter areas* (number of significant voxels in the left and right sides of each anatomical area)	x	y	z	T score	Corrected	Cluster	r**
					P-value (FWE)	size (voxel)	
<i>Positive correlation between 0-back activity and digit span score</i>							
Superior frontal gyrus (medial part) (L:2, R:11)	6	30	51	4.63	0.011	13	0.141
<i>Positive correlation between (2-back–0-back) activity and digit span score</i>							
Hippocampus (R:7)/Parahippocampal gyrus (R:5)	24	–15	–21	4.81	0.007	13	0.159
Postcentral gyrus (L:15)	–30	–30	75	4.80	0.008	17	0.121
Rolandic operculum (R:8)	51	–12	15	4.58	0.017	8	0.139
Parahippocampal gyrus (L:1)	–18	–3	–24	4.36	0.040	1	0.161

● *Labelings of the anatomical regions of gray matter were based on the WFU PickAtlas Tool (<http://www.fmri.wfubmc.edu/cms/software#PickAtlas/>) (Maldjian et al., 2004; Maldjian et al., 2003) and on 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.

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

Figure

Fig. 1

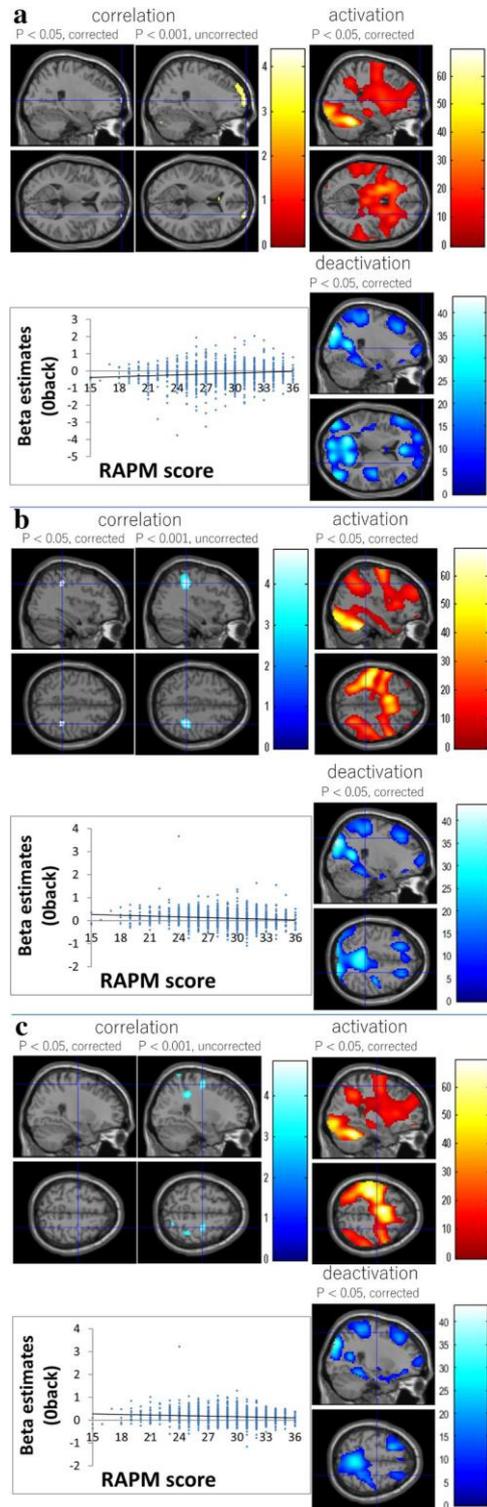


Fig. 2

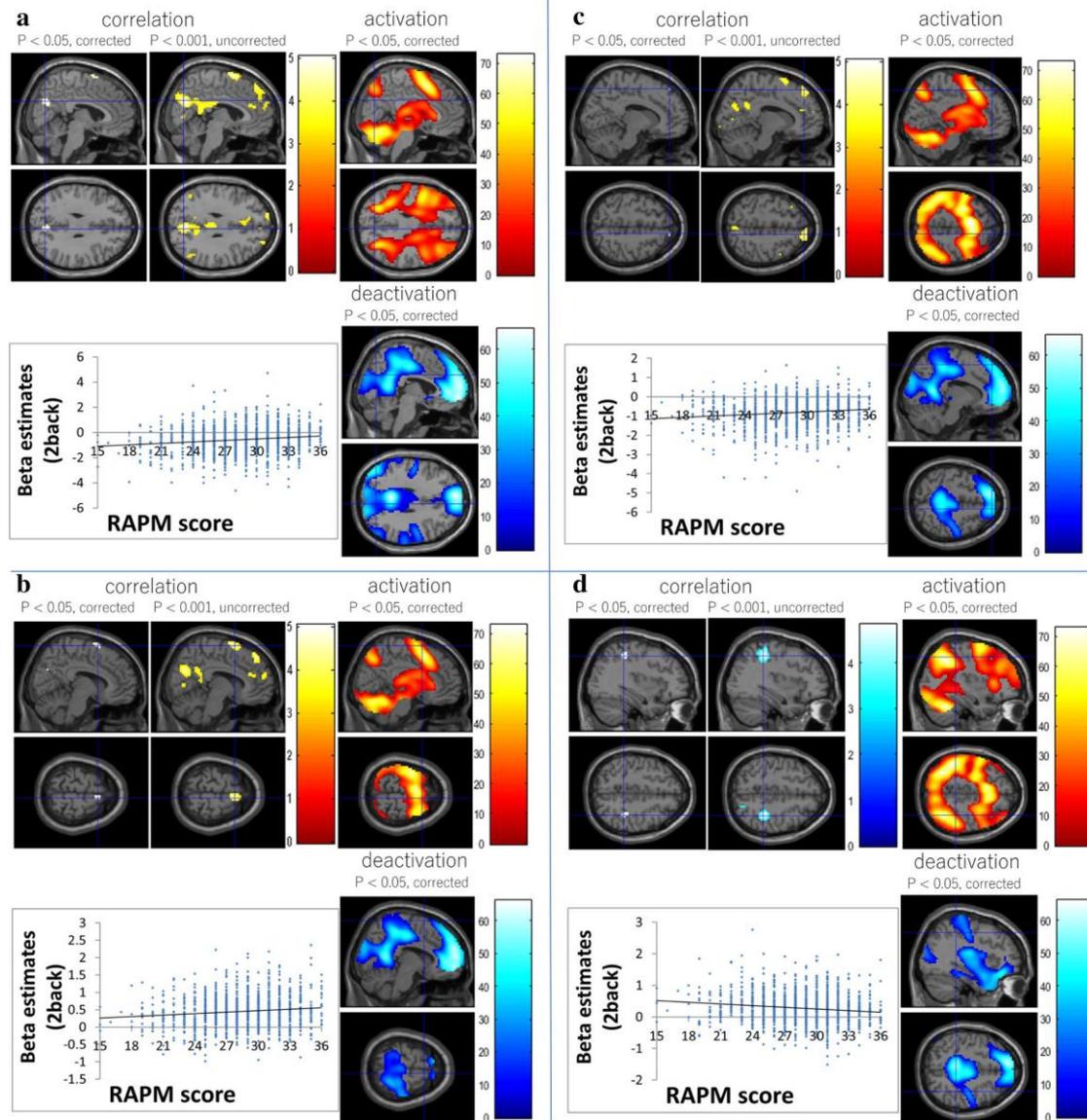


Fig. 3

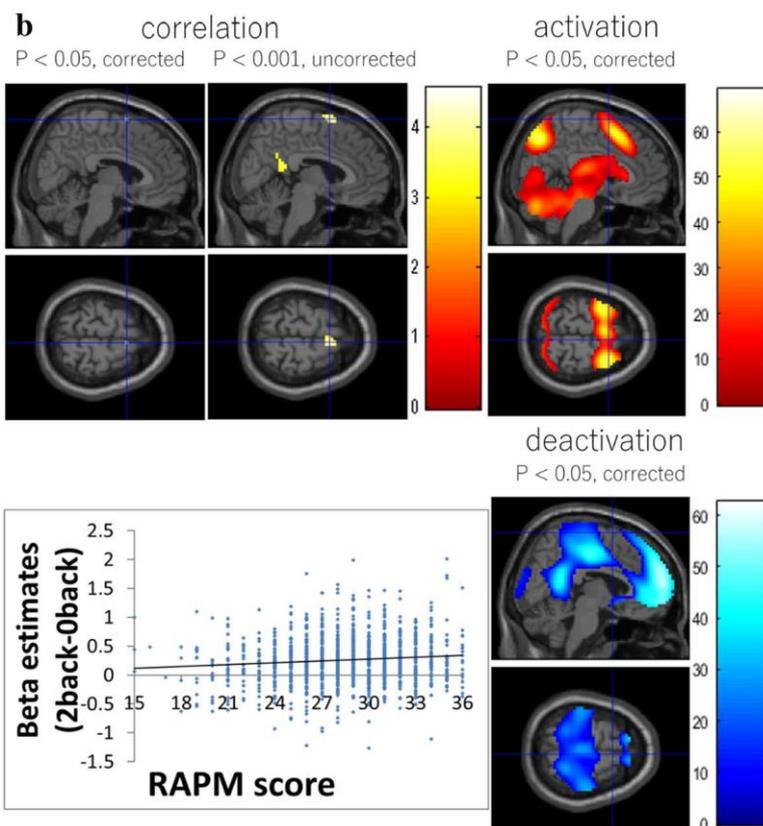
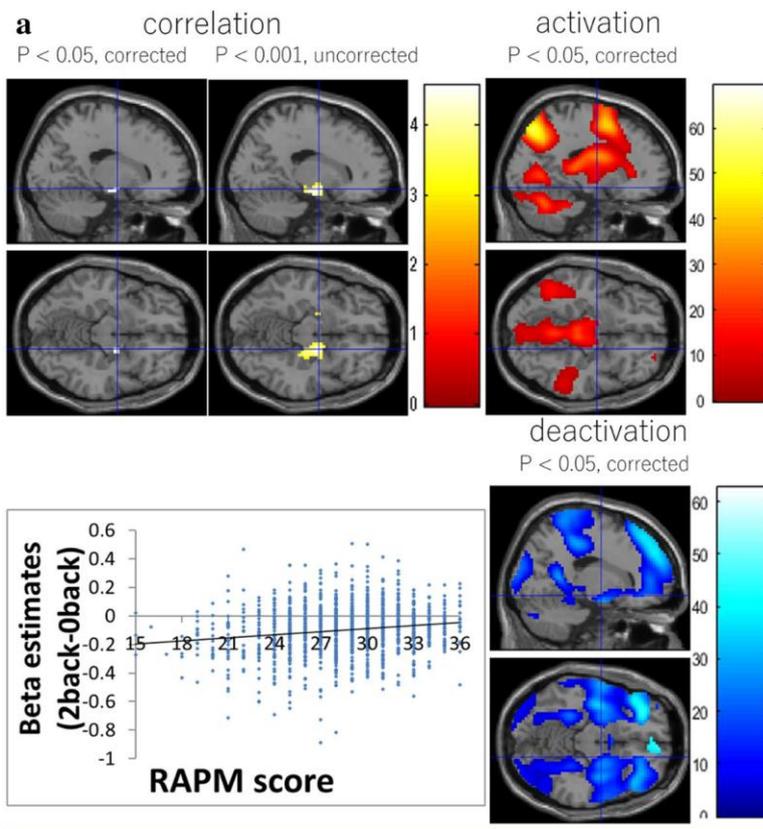


Fig. 4

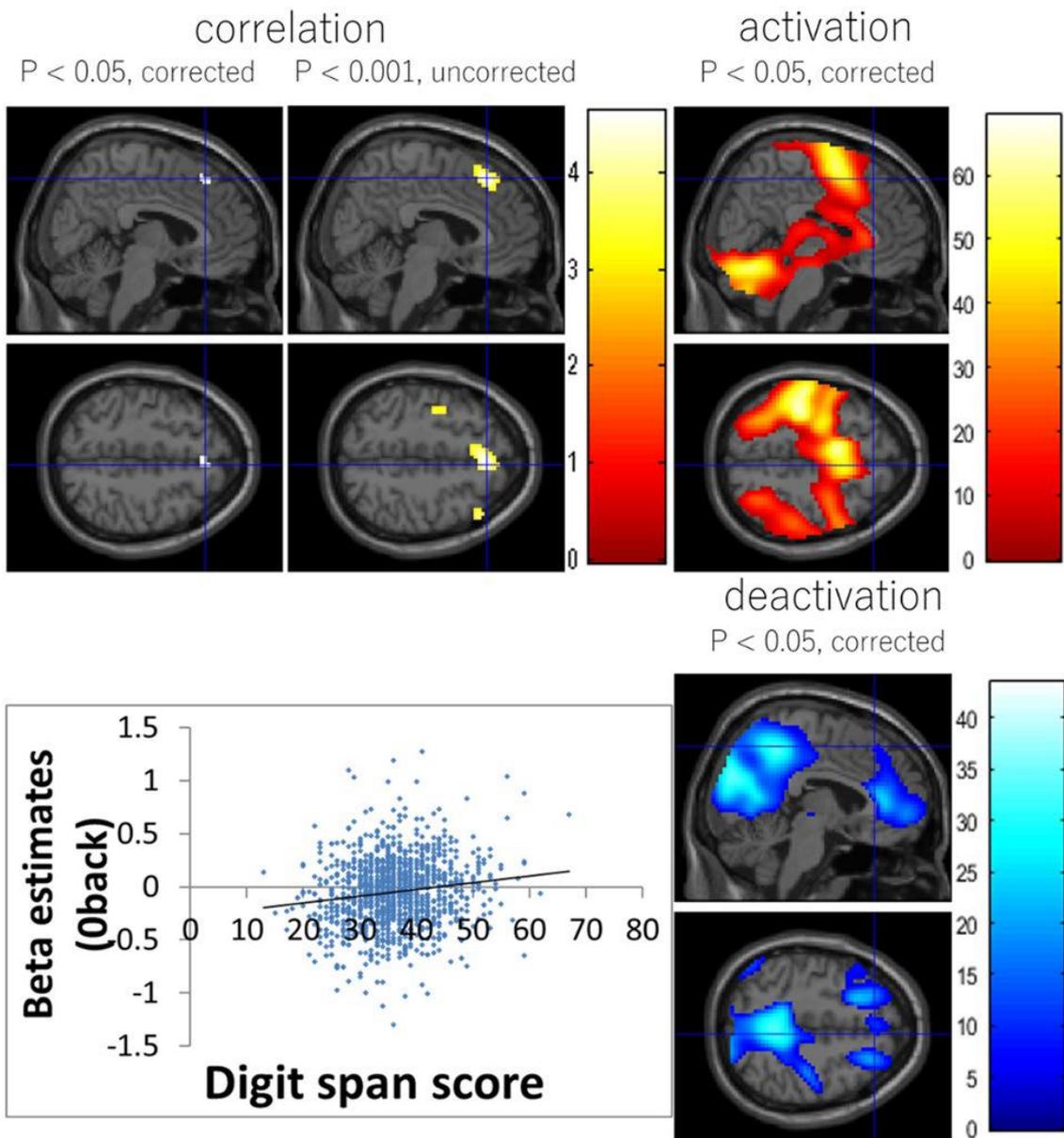


Fig. 5

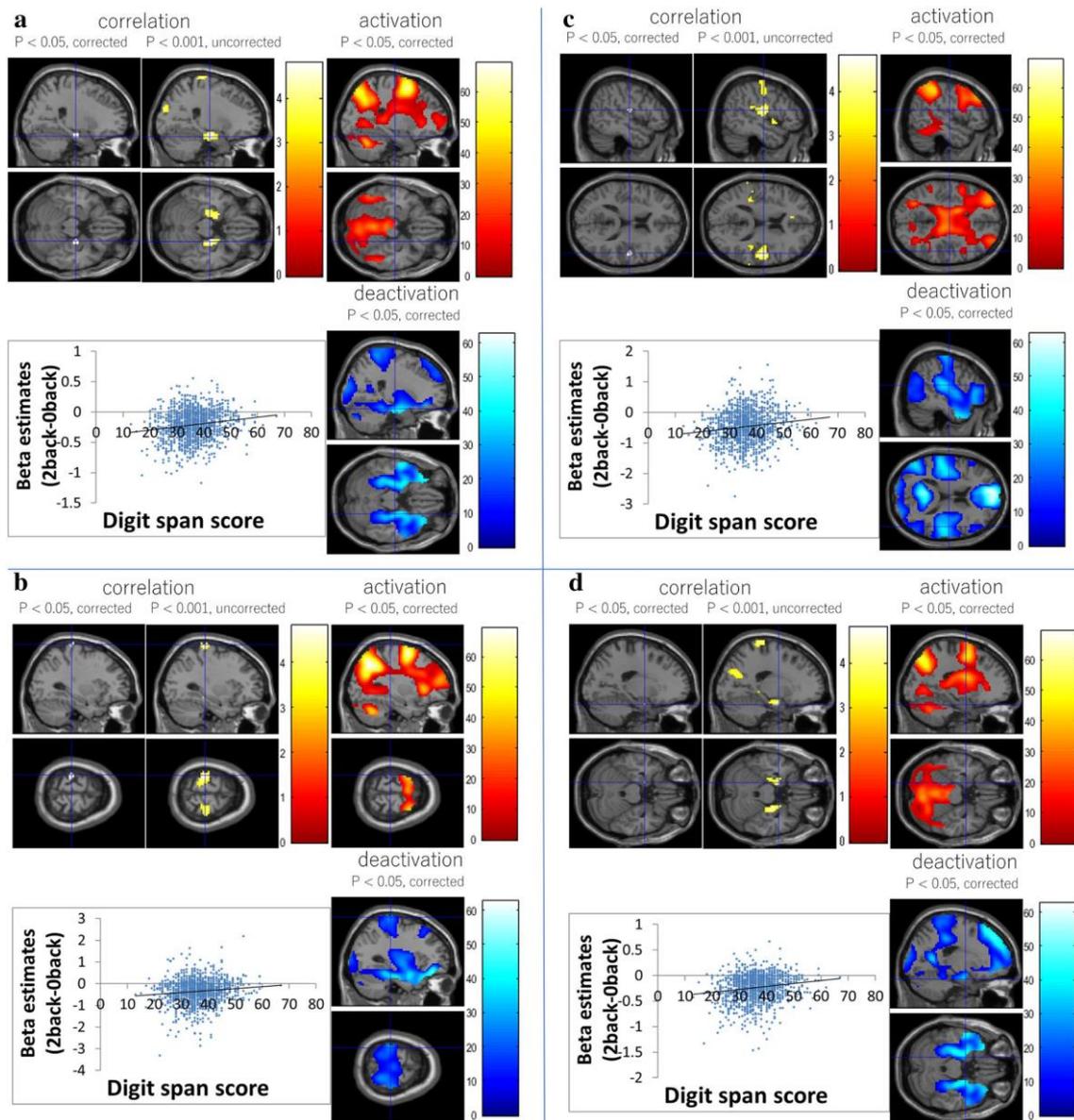


Figure legends

Fig. 1 Associations between regional neural activity during a simple cognitive task and performance of a general intelligence task. RAPM score and neural activity during the 0-back task showed a significant a positive correlation in the right RLPFC, b negative correlation in the right supramarginal gyrus, and c negative correlation in the right premotor cortex. In a–c, the colored bars indicate T scores. All results are overlaid on a “single-subject T1” SPM8 image. The scatter plots express the associations between psychometric scores and mean beta estimates of significant clusters. Upper left: areas of significant correlations. In the left column, the results were obtained using a threshold of $P < 0.05$, corrected for multiple comparisons based on 5000 permutations using the T score (subsequently termed ‘corrected’). In the right column, the results were obtained using a threshold of $P < 0.001$, uncorrected. Lower left: scatter plot of the association between RAPM score and mean beta estimates in the significant clusters. Upper right: areas of significant activation during the 0-back task. The results shown were obtained using a threshold of $P < 0.05$, corrected. Lower right: areas of significant deactivation during the 0-back task. Results were obtained using a threshold of $P < 0.05$, corrected

Fig. 2 Association between brain activity during the working memory task and general intelligence task score. RAPM score and brain activity during the 2-back task showed a significant a positive correlation in the precuneus; b positive correlation in the pre-supplementary motor area; c positive correlation in the dorsomedial prefrontal cortex, and d negative correlation in the right intraparietal sulcus. In a–d, the colored bars indicate T scores. All results are overlaid on a ‘single-subject T1’ SPM8 image. The scatter plots express the associations between psychometric scores and mean beta

estimates of significant clusters. Upper left: areas of significant correlations. In the left column, the results were obtained using a threshold of $P < 0.05$, corrected. In the right column, the results shown were obtained using a threshold of $P < 0.001$, uncorrected. Lower left: scatter plot of the association between RAPM score and mean beta estimates in the significant clusters. Upper right: areas of significant activation during the 2-back task. Results were obtained using a threshold of $P < 0.05$, corrected. Lower right: areas of significant deactivation during the 2-back task. Results were obtained using a threshold of $P < 0.05$, corrected

Fig. 3 Associations between working memory-specific activity and performance of a general intelligence task. The RAPM score and working memory-specific brain activity (2-back task– 0-back task contrast) showed a significant a positive correlation in the right hippocampus and b positive correlation in the right pre-supplementary motor area. In a and b, colored bars represent T scores. Results are overlaid on a “single-subject T1” SPM8 image. The scatter plots expressed the associations between the psychometric scores and mean beta estimates of significant clusters. Upper left: areas of significant correlations. In the left column, the results were obtained using a threshold of $P < 0.05$, corrected. In the right column, the results were obtained using a threshold of $P < 0.001$, uncorrected. Lower left: scatter plot of the association between RAPM score and mean beta estimates in the significant clusters. Upper right: areas of significant activation for the 2-back–0-back contrast. Results were obtained using a threshold of $P < 0.05$, corrected. Lower right: areas of significant deactivation for the 2-back–0-back contrast. Results were obtained using a threshold of $P < 0.05$, corrected

Fig. 4 Association between activity during a simple cognitive task and performance of a working memory span task. Digit span score and neural activity during the 0-back task showed a significant positive correlation in the dorsomedial prefrontal cortex. Colored bars represent T scores. Results are overlaid on a “single-subject T1” SPM8 image. The scatter plots express the associations between the psychometric scores and mean beta estimates of significant clusters. Upper left: areas of significant correlations. In the left column, results were obtained using a threshold of $P < 0.05$, corrected. In the right column, the results were obtained using a threshold of $P < 0.001$, uncorrected. Lower left: scatter plot of the association between digit span score and mean beta estimates in the significant clusters. Upper right: areas of significant activation during the 0-back task. Results were obtained using a threshold of $P < 0.05$, corrected. Lower right: areas of significant deactivation during the 0-back task. Results were obtained using a threshold of $P < 0.05$, corrected

Fig. 5 Associations between working memory-specific brain activity and performance of a working memory span task. The digit span score and working memory-specific brain activity (2-back task– 0-back task) showed significant positive correlations in a right hippocampus, b right postcentral gyrus, c right Rolandic operculum, and d left hippocampus. In a–d, colored bars represent T scores. All results are overlaid on a “single-subject T1” SPM8 image. The scatter plots express the associations between the psychometric scores and mean beta estimates of significant clusters. Upper left: areas of significant correlations. In the left column, the results were obtained using a threshold of $P < 0.05$, corrected. In the right column, the results were obtained using a threshold of $P < 0.001$, uncorrected. Lower left: scatter plot of the association between digit span

score and mean beta estimates in the significant clusters. Upper right: areas of significant activation for the 2-back–0-back contrast. Results were obtained using a threshold of $P < 0.05$, corrected. Lower right: areas of significant deactivation for the 2-back–0-back contrast. Results were obtained using a threshold of $P < 0.05$, corrected

Supplemental Online Material

Supplemental Methods

Subjects. The present study, which is part of an ongoing project to investigate the association between brain imaging, cognitive function, and aging, included relevant cognitive measures and imaging data from 1235 healthy, right-handed individuals (710 men and 525 women). The mean age (\pm standard deviation, SD) of the subjects was 20.8 ± 1.8 years (range, 18–27 years). Some of the subjects who participated in this study also became subjects in our intervention studies (psychological data and imaging data recorded before the intervention was used in this study, so this point is not supposed to affect the results) (Takeuchi et al., 2014). Psychological tests and MRI scans not described in this study were performed together with those described in this study. 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. For nonadult 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 foods and drinks on the day of cognitive tests and MRI scans. 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 using the exactly same methods (Takeuchi et al., 2015a).

Details of N-back fMRI task. Participants received instructions and practiced the tasks before entering the scanner. During scanning, they viewed stimuli on a screen via a mirror mounted on a head coil. Visual stimuli were presented using Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA). A fiber-optic light-sensitive key press interface with a button box was used to record participants' task responses.

Two conditions were used: 0-back and 2-back. Each condition had six blocks, and all N-back tasks were performed in one session. Subjects were instructed to recall visually presented stimuli (four Japanese

vowels) presented “n” letters before that currently presented (e.g., two letters previously for the 2-back task and the currently presented letter for the 0-back task). Two buttons were used during the 0-back task, and subjects were asked to push the first button when defined target stimuli were presented and the second button when non-target stimuli were presented. During the 2-back task, subjects were asked to push the first button when the currently presented stimulus and the stimulus presented two letters previously were the same and to push the second button when the currently presented stimulus and the stimulus presented two letters previously were different. Since the four stimuli were presented randomly, the ratio of matched trials to unmatched trials was 1:3 on average. Our version of the N-back task was designed to require individuals to push buttons continuously during the task period. The task level of the memory load was shown above the stimuli for 2 s before the task started, and remained visible during the task period (cue phase) and remained unchanged during the task period. Each letter was presented for 0.5 s, and a fixation cross was presented for 1.5 s between each item. Each block consisted of 10 stimuli. Thus, each block lasted 20 s. A baseline fixation cross was presented for 13 s between the last task item and the presentation of the next task level of the memory load (start of the cue phase). Thus, the rest period lasted for 15 s (13 s + 2 s). There were six blocks for each 2- and 0-back condition.

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, which was then realigned to the mean $b = 0$ image as described previously (Takeuchi et al., 2011a). 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 \times 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. (2011a) 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 $1.5 \times 1.5 \times 1.5 \text{ mm}^3$ 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 $1.5 \times 1.5 \times 1.5 \text{ mm}^3$. The voxel size of normalized BOLD images was $3 \times 3 \times 3 \text{ mm}^3$. 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.

Supplemental Discussion

Potential limitations of the study

This study has 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., 2011b; Takeuchi et al., 2010a, b; Wei et al., 2013) as their unique sample characteristics may bias results and limit generalizability. For example, young subjects with greater cognitive ability showed less deactivation in the TNN, a pattern also observed in elderly subjects with cognitive decline and patients with schizophrenia (Sambataro et al., 2008; Whitfield-Gabrieli et al., 2009). Whether our findings would also hold across the full range of population samples must be determined in future studies (Takeuchi et al., 2012). In addition, there are several discrepancies with previous studies. However, many of these previous studies have statistical problems (see Introduction). These studies also differ from ours in sample characteristics, task characteristics, and methodologies (Table 1), which may contribute to the discrepancies in results. Another possible limitation of the study is that we used only one measure for general intelligence (RAPM). However, the RAPM is one of the most broadly accepted single measures for general intelligence and has been frequently used alone (Raven, 1998). Also, if we used several time-consuming measures of general intelligence, it would have been impossible to collect the huge sample size needed to achieve our study goals. Nonetheless, the use of other types of general intelligence tests requiring verbal processes may increase the generalizability of our findings.

References

- 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.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97-113.
- 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.
- Sambataro, F., Murty, V.P., Callicott, J.H., Tan, H.Y., Das, S., Weinberger, D.R., Mattay, V.S., 2008. Age-related alterations in default mode network: Impact on working memory performance. *Neurobiology of Aging*.
- 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.
- Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., Kawashima, R., 2011a. 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., 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., 2012. The Association between Resting Functional Connectivity and Creativity. *Cerebral Cortex* Epub ahead of print.
- Takeuchi, H., Taki, Y., Nouchi, R., Hashizume, H., Sekiguchi, A., Kotozaki, Y., Nakagawa, S., Miyauchi, C.M., Sassa, Y., Kawashima, R., 2014. Effects of Multitasking-

Training on Gray Matter Structure and Resting State Neural Mechanisms. *Human Brain Mapping* 35, 3646-3660.

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., 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., 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.

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.

Whitfield-Gabrieli, S., Thermenos, H.W., Milanovic, S., Tsuang, M.T., Faraone, S.V., McCarley, R.W., Shenton, M.E., Green, A.I., Nieto-Castanon, A., LaViolette, P., 2009.

Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America* 106, 1279-1284.

Supplemental Table 1.

Demographics of study participants

Measure	Number			
	Mean	SD	Men	Women
Age	20.81	1.89	711	525
RAPM	28.45	3.83	711	525
Digit span	36.20	6.97	708	522
0-back-RT (ms)	4541.10	761.16	711	525
0-back-accuracy (%)	99.57	2.22	711	525
2-back-RT (ms)	6732.78	1789.77	711	525
2-back-accuracy (%)	99.2	2.84	711	525
Volume-wise frame-wise displacement*	0.21	0.05	711	525

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