博士論文（要約）

Research on functional roles of the inferior frontal gyrus using linguistic and numerical tasks

（言語課題と数課題を用いた下前頭回の
機能的役割の研究）

中井　智也

Tomoya Nakai

Department of Life Sciences, Graduate School of Arts and Sciences,

The University of Tokyo

Table of Contents

ii

List of Tables

List of Figures

List of Abbreviations

Abbreviations	Meanings
ACC	anterior cingulate cortex
AG	angular gyrus
ATL	anterior temporal lobe
BA	Brodmann's area
BOLD	blood-oxygen-level dependent
CFG	context free grammar
CFT	category fluency task
Cr	creatine
(MB-)EPI	(multi-band) echo planar imaging
FA	flip angle
FDR	false-discovery rate
FG	fugiform gyrus
(f)MRI	(functional) magnetic resonance imaging
FOV	field of view
GABA	gamma-aminobutyric acid
GcSS	group-constrained subject-specific
GLM	general linear model
H student	student in humanity/social science departments
IFG	inferior frontal gyrus
IFGoper	opercular part of the IFG
IFGtri	triangular part of the IFG
IPL (IPS)	intraparietal lobule (sulcus)
L	left hemisphere
Lang	linguistic task
LB	left-branching
LFT	letter fluency task
LOG	lateral occipital gyrus
LQ	laterality quotient
M	medial
Num	numerical task
MCC	middle cingulate cortex
MEGA-PRESS	Mescher-Garwood point-resolved spectroscopy
MEG	magnetoencephalography

MNI	Montreal neurological institute
MSE	mean square error
MTG	middle temporal gyrus
MOG	middle occipital gyrus
(H^1-)MRS	proton-magnetic resonance spectroscopy
MRRAGE	magnetization-prepared rapid acquisition gradient-echo sequence
N student	student in natural science department
OP	occipital pole
OTG	occipitotemporal gyrus
PCC	poster cingulate cortex
PCG	posterior central gyrus
PET	positron emission tomography
PG	parahippocampal gyrus
R	right hemisphere
rANOVA	analysis of variance with repeated measures
RB	right-branching
ROI	region of interest
RS	repetition suppression
RT	reaction time
SA	sense of accomplishment
SD	standard deviation
SEM	standard error of the mean
SMA	supplementary motor area
SPM	statistical parametric mapping
STG	superior temporal gyrus
TE	echo time
TR	repetition time
VBM	voxel-based morphometry
vlPFC	ventrolateral prefrontal cortex
vmPFC	ventromedial prefrontal cortex
VOI	volume of interest
WM	working memory

Chapter 1. General Introduction

1.1. Language processing in the inferior frontal gyrus

Language is one of the cognitive functions which was reported for the localization of brain function for the first time in the history of neuroscience. In the 1861, Pierre Paul Broca reported that a patient with a left-hemisphere damage had difficulty in language production, and could only say "tan, tan" (Broca, 1861a, 1861b) (but see Prins & Bastiaanse (2006) for the prehistory of aphasia reports). The critical region of this patient Leborgne (and later with another patient Lelong) located in the left inferior frontal gyrus (IFG) is called as Broca's area, although modern neuroimaging technology revealed that the lesion of these patients extended to the large medial regions and had a damage on the superior longitudinal fasciculus (Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007). After Broca's discovery, many neuropsychological evidence of language impairment has been accumulated (Dejerine, 1891; Exner, 1881; Wernicke, 1874), leading to the Geschwind model which assigned a function of language production to the IFG (Geschwind, 1965, 1970).

Recent development of neuroimaging techniques provide evidence against the traditional Geschwind model of language areas. With Positron emission tomography (PET), Stromswold et al. (1996) reported the relation of the IFG for the syntactic processing for the first time, by comparing center-embedded and right-branching sentences. A functional magnetic resonance imaging (fMRI) study suggest that, in the subparts of the IFG, syntactic processing is related to the BA44/45, but semantic processing of sentences is related to BA47 (Dapretto & Bookheimer, 1999). Researchers in neurobiology of syntax gradually took the concepts in modern linguistic theories into neuroscience. Neuroimaging studies reported that processing of strings based on the hierarchical dependency rules induced activation in the left IFG (Bahlmann, Schubotz, & Friederici, 2008; Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006), although their task design still had problem for fully capturing hierarchical dependency (Ojima & Okanoya, 2014). Using sentence stimuli with various syntactic structures, Pallier, Devauchelle, and Dehaene (2011) showed that activation in the left IFG increased with syntactic constituent size, even with jabberwocky sentences without

semantic components. These studies suggest the critical involvement of the left IFG in syntactic processing in language.

1.2. Anatomical feature of Inferior frontal gyrus

The IFG roughly corresponds to Brodmann's area (BA) 44 (pars opercularis) and 45 (pars triangularis). Here I dente BA 44 as IFGoper, and BA 45 as IFGtri. This naming is based on the cytoarchitectonic study of Brodmann in 1909 (Zilles & Amunts, 2010). The major problem of Brodmann's mapping and following anatomical labelling studies were that they generalize the brain labelling from a few brain data. Indeed, Amunts et al. (1999) developed techniques to quantify laminar distributions of cell densities in the post-mortem human brain, and reported large individual variability of cytoarchitectonic border of the IFGoper and IFGtri. This study also reported a left-ward lateralization in the IFGoper, but not in the IFGtri. The left-ward lateralization of the IFG is also reported in other studies which used voxel-based morphometry (VBM) (Foundas, Eure, Luevano, & Weinberger, 1998; Foundas, Leonard, Gilmore, Fennell, & Heilman, 1996; Uylings, Jacobsen, Zilles, & Amunts, 2006).

According to Zilles & Amunts (2010), the most important and distinguishable aspect of Brodmann's work was that he introduced an evolutionary perspective in his anatomical labelling. Comparison of the cytoarchitecture of human and macaque monkey indicate the correspondent part of the IFGoper an IFGtri in monkey's frontal cortex (Petrides & Pandya, 2002), in that the layer III of the IFGtri contains large and deeply stained pyramidal neurons, while the IFGoper has barely-distinguishable layer IV. Even the leftward lateralization was found for monkeys (Cantalupo & Hopkins, 2001; Hopkins et al., 2009), but the detailed lateralization property in the IFGoper/IFGtri is not observed yet (Schenker et al., 2010). So far, although we observe relatively huge brain size of Homo sapiens (Dunbar, 1993), the accumulated results make it difficult to provide a clear border between human and other primates in terms of localized anatomical features.

1.3. Neurobiology of mathematics

It has been proposed that estimation of numerical quantity and serial counting is based on the different systems (Izard, Pica, Spelke, & Dehaene, 2008). In addition to the

inferior parietal lobule　(IPL), serial counting may recruit the left premotor cortex (Kansaku et al., 2006; Piazza, Mechelli, Price, & Butterworth, 2006). It seems that approximate arithmetic and exact arithmetic may recruit distinct cortical areas. The former activated the bilateral inferior parietal sulcus (IPS), while the latter induced left-lateralized activations including left prefrontal cortex (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Ischebeck, Zamarian, Egger, Schocke, & Delazer, 2007; Menon, Mackenzie, Rivera, & Reiss, 2002). Activations in the left frontal area were also found in abstract algebra (Qin et al., 2004), or integral calculus (Krueger et al., 2008). The individual variability of connectivity property between left IPL and left IFG can predict arithmetic skills of children (Tsang, Dougherty, Deutsch, Wandell, & Ben-Shachar, 2009). In summary, human-specific capacity of counting and calculation with exact quantity appears to recruit additional brain regions in the left frontal cortex.

1.4. Shared neural basis for language and mathematics?

There has been a controversy on the shared neural bases for mathematics and language (Baldo & Dronkers, 2007; Klessinger, Szczerbinski, & Varley, 2007; Makuuchi, Bahlmann, & Friederici, 2012; Maruyama et al., 2012; Monti, Parsons, & Osherson, 2012; Nakai & Sakai, 2014; Varley, Klessinger, Romanowski, & Siegal, 2005). Most of the previous studies on this issue concerns the similarity of syntactic (grammatical) property in both domains. For example, Varley et al. (2005) showed that even patients with large left-hemisphere perisylvian lesions and severe grammatical impairment showed various numerical capacities, indicating the independence of the language ability and mathematical ability. In contrast, a close association between acalculia and aphasia has been suggested in the several neuropsychological reports (Basso, Burgio, & Caporali, 2000; Basso, Caporali, & Faglioni, 2005; Delazer, Girelli, Semenza, & Denes, 1999). Moreover, another lesion study with a large number of patients suggest that damage to the left IFG affect for both aphasic test scores and calculation scores (Baldo & Dronkers, 2007).

In the fMRI and MEG study of Maruyama et al. (2012), the authors asked participants to passively view two mathematical expressions (m-expressions) and detect changes in visual features (e.g., the operator swapping between "+" and "−") without any calculation, and reported small structure-sensitive activations in the visual areas, but

not in the language areas. In contrast, Makuuchi et al. (2012) used a calculation task with the Reverse Polish notation (e.g., [[3 2 –] [2 4 +] ×]) and a sentence comprehension task with structurally matched sentences, and found joint activation in syntactic processing for linguistic and numerical tasks in the left IFG, which has been considered as a central region for language processing. Moreover, the fMRI study of Nakai and Sakai (2014) revealed that activation in the left IFG was modulated by the structural complexity of calculations. These observations suggest that there should be a partially shared system for linguistic and numerical processing based on their structural information in the human brain, and that share system may be anatomically located in the left IFG.

The major disagreements on this issue are as follows: Firstly, working memory is also a candidate for explain the role of IFG. A previous study argued that working memory required in calculation induced activations in the left IFG for a numerical task (Ischebeck, Zamarian, Egger, Schocke, & Delazer, 2007). Secondly, multiple cognitive domains are associated with activations in the left IFG. Even if two different domains have activation overlaps in the left IFG, there is a possibility that different neural circuits exist in the same region. Indeed, a previous study using subject-specific method showed different activation patterns of multiple cognitive domains in the left IFG (Fedorenko, Duncan, & Kanwisher, 2012). Thirdly, syntactic structures in m-expressions were confounded with the effect of operators. Hung et al. (2015) claimed that in the previous study of Nakai and Sakai (2014) the number of operators (addition and subtraction) also increased with m-expressions with complex syntactic structures. The idea of the shared neural basis (i.e., the left IFG) for the linguistic and numerical tasks is not supported unless these problems are answered. Here I prepared three experiments to answer these questions. In the experiment 2 & 3 I compared exactly the same set of numerical stimuli with the same working memory load (but with different combination with the prime linguistic stimuli). Therefore activations are independent of working memory. In the experiment 3 I further excluded the possibility that the linguistic and numerical tasks activated the same region by chance, by using repetition suppression (RS) effect. The RS effect suggests that two stimuli are processed in the same neural circuit. In the experiment 4 I prepared task with both number and letter

notations. Neural activation induced by letter notation cannot be explained by the number of operators.

1.5. Theoretical background

Abstract syntactic structures can be recognised in both language and mathematics. An m-expression such as "$8 + 4 \times 3$" is composed first by multiplication of 4 with 3, and then the result of the multiplication is combined with 8 by addition ($[8 + [4 \times 3]]$). Similarly, a sentence such as "John ate apples" is composed first by the combination of "ate" and "apples", and then the resultant phrase is combined with "John" to make a whole sentence ([John [ate apples]]). According to the generative syntax theory, syntactic structure in natural language is based on the recursive computation of words and phrases (Chomsky, 1957). The ability of recursive computation, or recursion, is considered to be a pivotal feature to distinguish human language from other communicative systems (Friederici, Bahlmann, Friedrich, & Makuuchi, 2011; Pinker & Jackendoff, 2005).

Theoretically, a symbolic system with recursive computation can be modelled by context-free grammar. The formal definition of the context free grammar (CFG) is described as follows (Hopcroft, Motwani, & Ullman, 2006):

Context free grammar G is a 4-tuple $G = (V, T, S, P)$ where:
- V is a finite set of non-terminal variables
- T is a finite set of terminal alphabets
- $S \in V$ is the start variable
- P is a finite set of production rules, where each production rule has the form $V \rightarrow (V \cup T)^*$.

Note that I do not provide mathematically rigorous definition, and omit some additional definitions in order to simplify the following discussion. M-expressions are often used as examples of the CFG description. Although there seems to be no widely-accepted description style (i.e., the notations of variables or production forms) of m-expressions,

here I provide a tentative CFG description of m-expressions necessary for the discussion in the current article. I assume $G = (V, T, S, P)$ where

$V = \{M, Num, Op, Eq\}$

$T = \{1,2,3, \ldots, +, -, \times, \div\}$

$S = M$

P is given as follows:

(1) $M \rightarrow Num, Num\ Eq\ Num$

(2) $Num \rightarrow Num\ Op\ Num$

(3) $Num \rightarrow 0,1,2, \ldots, 9$

(4) $Op \rightarrow +, -, \times, \div$

(5) $Eq \rightarrow =$

For example, an m-expression "$1 + 2 \times 3 = 7$" is derived as follows:

$M \rightarrow Num\ Eq\ Num$

$\rightarrow (Num\ Op\ Num)\ Eq\ Num$

$\rightarrow \big(NumOp\ (Num\ Op\ Num)\big)\ Eq\ Num$

$\rightarrow \big(\ 1\ +\ (\ 3\ \times\ 2\ )\big) =\ 7$

Note that I used parentheses for the sake of intelligibility. The derivation process becomes visually more intelligible if a *tree diagram* (or *tree structure*) is used (Figure 1).
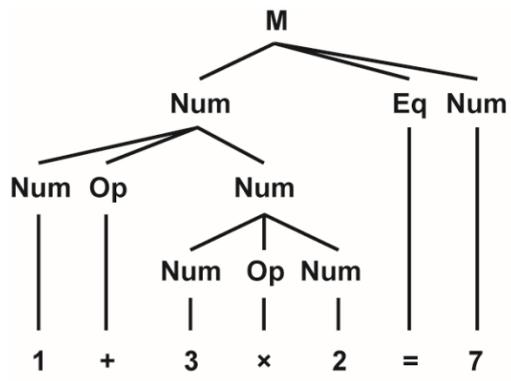
Figure 1. A hierarchical tree structure of the m-expression.

Such tree structure is *hierarchical* if derivation (i.e., production) is applied more than once. In other word, the derivation (or computation) is applied *recursively*. Therefore, it is identical to say that the production process is based on the "recursive computation" and that the product has a "hierarchical tree structure". In some cases operators are not included as constituents in the tree structure (Maruyama, Pallier, Jobert, Sigman, & Dehaene, 2012; Nakai & Sakai, 2014), and I also omitted operators in the tree structure in Chapter 3. This description, however, does not change the hierarchical feature of tree structure.

Noam Chomsky applied the CFG to sentences of natural language for the first time (Chomsky, 1957). Although Chomsky further introduced transformation rules to explain linguistic phenomena such as a semantic correspondence of active and passive sentences, there are other theories of grammars which do not include transformational rules [e.g., Combinatorial Categorical Grammar (Steedman, 2000)]. Since natural language seems to have structures which cannot be derived by CFG, natural language may have "mildly context-sensitive grammar" (Joshi, 1985). However, it is clear that most of the sentences in natural language can be described by CFG, thus as a start point I focus on the sentences which can be described by CFG.

According to the above background, it has been argued that the capacity of recursive computation provides a basis not only for syntactic processing in language but also for mathematics including the natural number system (Chomsky, 2006; Devlin, 2000; Hurford, 1987). Indeed, there have been attempts to describe syntactic structures of m-expressions (Ernest, 1987). Such shared computational basis of language and mathematics leads to the idea that mathematics has its origin in pre-existing linguistic capacity (Devlin, 2000). Therefore, whether the syntactic processing is actually shared in language and mathematics is a core research question in considering the origin of human-unique capacity.

1.6. Measuring neurotransmitters in the IFG

As discussed in the above sections, there has been a lot of studies which focused on the functional role of the IFG using linguistic or numerical tasks, as well as studies on anatomical features in the IFG. No study, however, has been reported on the neurotransmitter basis of this region. Recent studies reported the relationship between various behavioural measurements and GABA concentration in corresponding brain

regions (Edden, Muthukumaraswamy, Freeman, & Singh, 2009; Boy et al., 2010; Puts, Edden, Evans, McGlone, & McGonigle, 2011; Terhune, Russo, Near, Stagg, & Cohen Kadosh, 2014). Measuring neurotransmitters in the IFG may provide additional information for this region, which cannot be acquired by the measurements of brain activations, and should thus be useful to clarify the specificity of the IFG from multiple points of view. In this study, I tried to measure the neurotransmitter basis of the IFG by introducing a technique of magnetic resonance spectroscopy (MRS).

Standard linguistic tasks have been developed for the sake of assessment of linguistic skills, and can be considered as a useful and reliable method for examining individual variability of linguistic skills. In this study I adopted the verbal fluency task (Oswa, Maeshima, Tanemura, Sekiguchi, & Itakura, 2006), which has been also used in the previous fMRI studies of the IFG functions (Costafreda et al., 2006). Compared to the domain of language which has a long history of neuroscientific evidence on the role of the IFG, the involvement of the IFG for the domain of mathematics is still controversial. I thus used only the linguistic task (verbal fluency task) in the experiment of MRS. A standard numerical task should be developed in future. In the current study I tried to clarify the functional overlap of linguistic and numerical tasks in the IFG, which would be useful for the further challenges of the development of numerical tasks appropriate for the use in future MRS studies.

1.7. Other brain regions necessary for language and mathematics

Human reward system would be also important in learning language and mathematics. So far, caudate nucleus (a part of the striatum) in the reward system is a putative region which plays a critical role in artificial grammar learning (Chan, Ryan, & Bever, 2013; Forkstam, Hagoort, Fernandez, Ingvar, & Petersson, 2006; Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004), while the involvement of reward system in mathematics is unclear. The recent study of white matter tracts (which are mainly composed of nerve fibers) with a diffusion tensor imaging technique suggested the direct anatomical connection between the left IFG and striatum (Ford et al., 2013). To understand the neural mechanism of language and mathematics, it is necessary to reveal functional integration of Broca's area and other brain regions such as reward system.

1.7. The aim of the current study

In the present series of experiments, I seek to reveal the functional specificity and neurotransmitter basis of the IFG using linguistic and numerical tasks. For this purpose I performed three experiments. In Experiment 1 (Chapter 2), I use MRS technique, and reveal neurotransmitter specificity of the IFG. In Experiment 2 (Chapter 3), I prepare linguistic and numerical stimuli with similar/dissimilar syntactic structures, and found cross-domain structural priming effect between two domains. In Experiment 3 (Chapter

4), I extend the experimental paradigm of Experiment 2 by using fMRI, and found cross-domain structural repetition suppression effect in the IFG. In Experiment 4 (Chapter 5), I extended the previous study of Nakai and Sakai (2014), and revealed the IFG activation with the recursive computation for both number and letter notations. The results of current experiments may contribute to understand fundamental features of the IFG.

Chapter 2. Neurotransmitter basis in the IFG

2.1. Introduction

After Broca's early report on a patient with speech impairment (Broca, 1861a, 1861b), a substantial amount of research and debate has focused on the function of the left IFG in language processing (see Grodzinsky & Santi (2008) for a review). Neuroimaging studies have reported activation in this region for phonological (Heim, Opitz, Müller, & Friederici, 2003), semantic (Newman & Twieg, 2001), and syntactic processing (Musso et al., 2003). Although the diverse results of those studies indicate multiple functions of the IFG, a common and particular feature of this region is its leftward functional/anatomical lateralization (Amunts, Schleicher, Ditterich, & Zilles, 2003). In contrast to such functional/anatomical differences, studies thus far have not shown a clear asymmetry of neurotransmitters in the cortical regions but, rather, individual variability in its asymmetrical distribution (Glick, Ross, & Hough, 1982; Rossor, Garrett, & Iversen, 1980).

Gamma-aminobutyric acid (GABA) is known as an inhibitory neurotransmitter in the adult human brain, and recent neuroimaging techniques have revealed strong correlations between GABA concentrations in the cortical regions and various behavioral measures, such as orientation discrimination (Edden, Muthukumaraswamy, Freeman, & Singh, 2009), motor control (Boy et al., 2010), tactile discrimination (Puts, Edden, Evans, McGlone, & McGonigle, 2011), and time perception (Terhune, Russo, Near, Stagg, & Cohen Kadosh, 2014). It has also been reported that resting-state GABA concentrations are correlated with blood-oxygen-level dependent (BOLD) signals in functional magnetic resonance imaging (fMRI) (Muthukumaraswamy, Edden, Jones, Swettenham, & Singh, 2009), suggesting the importance of GABAergic inhibition in task-related neural activation.

As for individual variability in the linguistic domain, verbal fluency tasks have traditionally been used for assessing dysfunction of linguistic ability for word production (Benton, 1968; Milner, 1964). Considering fMRI activation in the IFG during verbal fluency tasks (Amunts et al., 2004; Phelps, Hyder, Blamire, & Shulman, 1997), it is possible that GABAergic inhibition also plays a critical role for such linguistic tasks in the IFG. In this experiment, I used proton magnetic resonance spectroscopy (H^1-MRS),

which is a non-invasive technique to quantify the concentration of neurotransmitters and metabolites, to measure resting-state GABA concentrations bilaterally in the IFG. To examine the effects of baseline references, I used both Creatine (Cr) and water (H_2O) as references for the GABA signal. I hypothesized that GABA concentrations in this region are related to the individual variability of verbal fluency scores, and I also examined whether there is an asymmetry of neurotransmitter between the left and right IFG.

2.2. Methods
2.2.1. Participants
Twenty-eight healthy, native Japanese speakers (all males, aged 18–22 years) participated in the current experiment. Twenty-seven participants were right-handed (laterality quotient (LQ): 57.9–100), while one participant was left-handed (LQ: -64.7), according to the Edinburgh inventory (Oldfield, 1971). Prior to their participation in the study, written informed consent was obtained from all participants. This experiment was approved by the Ethics Committee of the University of Tokyo, Komaba.

2.2.2. Verbal Fluency Task
The verbal fluency task was composed of a category fluency task (CFT) and a letter fluency task (LFT). For the CFT, participants were given one minute to write down as many nouns as possible belonging to a certain category. I chose three categories (Animal, Fruit, and Vehicle) according to a previous study (Oswa, Maeshima, Tanemura, Sekiguchi, & Itakura, 2006). For the LFT, the participants were given one minute to write down as many nouns as possible that start with a certain letter. I chose three Japanese letters ("Shi", "I", "Re"). I evaluated the CFT and LFT scores by averaging the number of written nouns in the three categories and letter groups, respectively.

2.2.3. MRI Data Acquisition
MRI scans were conducted on a 3.0 T scanner (MAGNETOM Prisma; Siemens, Erlangen, Germany) with a 64-channel head coil. For anatomical reference, high-resolution T1-weighted images of the whole brain (176 sagittal slices, $1 \times 1 \times 1$ mm^3) were also acquired from all participants with a Magnetization-prepared Rapid Acquisition Gradient-echo

sequence [MPRAGE, repetition time (TR) = 2000 ms, echo time (TE) = 2.9 ms, flip angle (FA) = 9°, field of view (FOV) = 256×256 mm^2].

GABA-edited MRS data were acquired from a $20 \times 20 \times 20$ mm^3 volume positioned on the bilateral IFG using the Mescher-Garwood point-resolved spectroscopy (MEGA-PRESS) method (TR = 2000 ms, TE = 68 ms, 64 averages) (Mescher, Merkle, Kirsch, Garwood, & Gruetter, 1998). For odd-numbered acquisitions, a frequency-selective editing pulse was applied to the GABA at 1.9 ppm (ON), while the same pulse was applied at 7.5 ppm in the even-numbered acquisitions (OFF). MRS data was first acquired in the right IFG, then another MRS data was acquired in the left IFG. The center of the volume of interest (VOI) in the right IFG was positioned at the voxel 5 mm above the superior margin of the third cerebral ventricle, 10 mm rightward from the center of the right hemisphere (1/2 of the distance between the midline and lateral margin of the brain and 1/3 of the distance between the anterior and posterior margin of the brain), which was determined with the T1-weighted images. I determined the above voxel locations according to a previous MRS study (Michels et al., 2012) so as to encompass the dorsal part of BA 44. I selected BA 44 (but not BA 45) because the previous studies reported joint activations of linguistic and numerical tasks in BA 44 (Makuuchi et al., 2012). Note that this voxel also contained a part of the dorsolateral prefrontal cortex. The left IFG VOI was positioned symmetrically from the right IFG VOI with respect to the midline. I also acquired unsuppressed water reference spectra for each VOI.

2.2.4. MRS data analyses
MEGA-PRESS data were analyzed by GANNET software on MATLAB (Edden, Puts, Harris, Barker, & Evans, 2014). Since GABA spectra were co-edited with those of macromolecules, I denoted the measured spectra as GABA+. The ON and OFF spectra were subtracted to provide a spectrum containing the GABA+ peak while simultaneously removing the overlapped Cr peak. The GABA+ peak was fitted with a single Gaussian curve with a linear baseline over a range of 2.79 and 3.55 ppm, and the Cr peak was fitted with a Lorentzian function over a range of 2.72 and 3.12 ppm. Cr concentration at 3 ppm from the OFF spectrum, as well as unsuppressed water concentration at 4.7 ppm, was used as a reference for GABA+. Spectra with relative Fit Error of GABA+ or GABA+/Cr values more than three times the standard deviation from the mean were excluded in the

subsequent analyses. Accordingly, spectra from one participant in the right IFG and four participants in the left IFG were excluded.

2.3. Results

I extracted MRS data from the bilateral IFG and quantified the GABA+/Cr ratio in each region (Figure 2). The average fit errors of the remaining participants were 14.0 ± 3.3 % (mean $\pm$ SD) in the left IFG and 12.6 ± 2.9 % in the right IFG. I compared GABA+/Cr values in the left IFG and right IFG for each participant and found no significant differences between them (left IFG, 0.114 ± 0.059; right IFG, 0.117 ± 0.057; $t(22)= 0.17$, $p = 0.86$). To show that the current results are independent of the baseline reference used, I also quantified GABA+/H_2O ratios for the same participants and again found no significant differences between the left IFG and right IFG (left IFG, 0.98 ± 0.55; right IFG, 0.96 ± 0.47; $t(22)= 0.16$, $p = 0.87$).
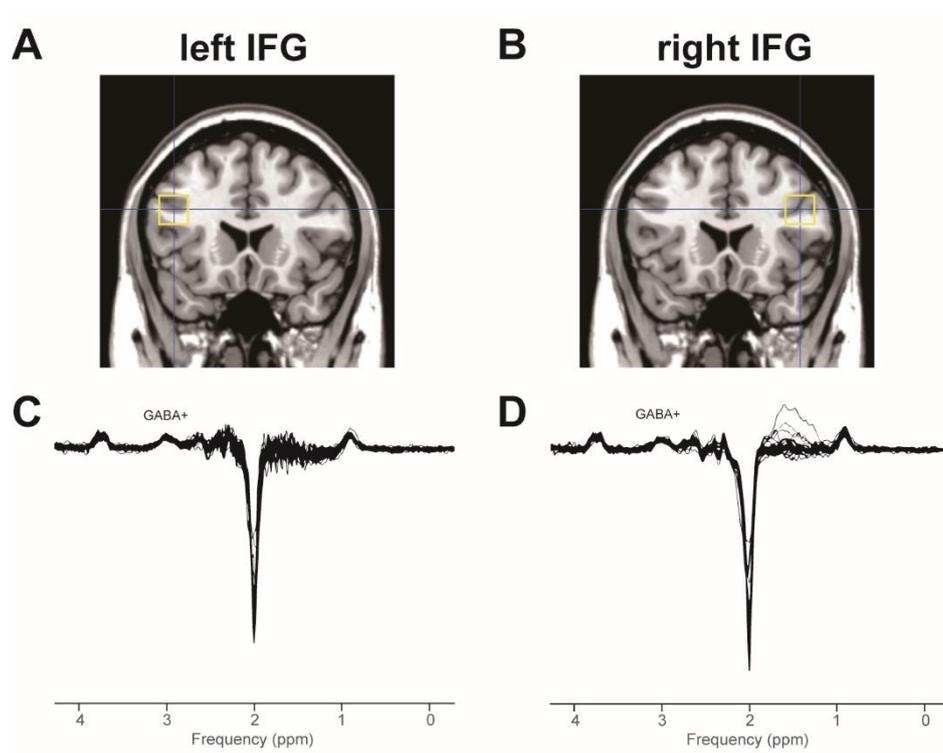
Figure 2.MRS voxels and spectra. Voxel locations of the bilateral IFG on the coronal surface (MNI coordinate, y = 20) of the standard brain (A, B). ON-OFF subtracted spectra from all participants are shown for both left IFG (C) and right IFG (D).

Next, I performed correlation analyses between GABA+/Cr values and verbal fluency test scores (CFT and LFT). I found a significant negative correlation between CFT scores and GABA+/Cr in the left IFG (Figure 3A, Pearson's correlation coefficient $r = -0.48$, $p = 0.017$) but not in the right IFG (Figure 3B, $r = -0.33$, $p = 0.094$). In both regions, I found no significant correlation between LFT scores and GABA+/Cr (Figure 3C, left IFG; $r = -0.30$, $p = 0.15$; Figure 3D, right IFG; $r = 0.05$, $p = 0.82$). Note that the correlation with the CFT scores in the left IFG was marginally significant with multiple comparisons (significance level $\alpha = 0.0125$, Bonferroni correction). This effect was robust even when I excluded one left-handed participant; there was a significant negative correlation between CFT scores and GABA+/Cr in the left IFG (left IFG; $r = -0.45$, $p < 0.032$; right IFG; $r = -0.29$, $p = 0.15$), but no correlation was found between LFT scores and GABA+/Cr (left IFG; $r = -0.24$, $p = 0.27$; right IFG; $r = 0.16$, $p = 0.44$).
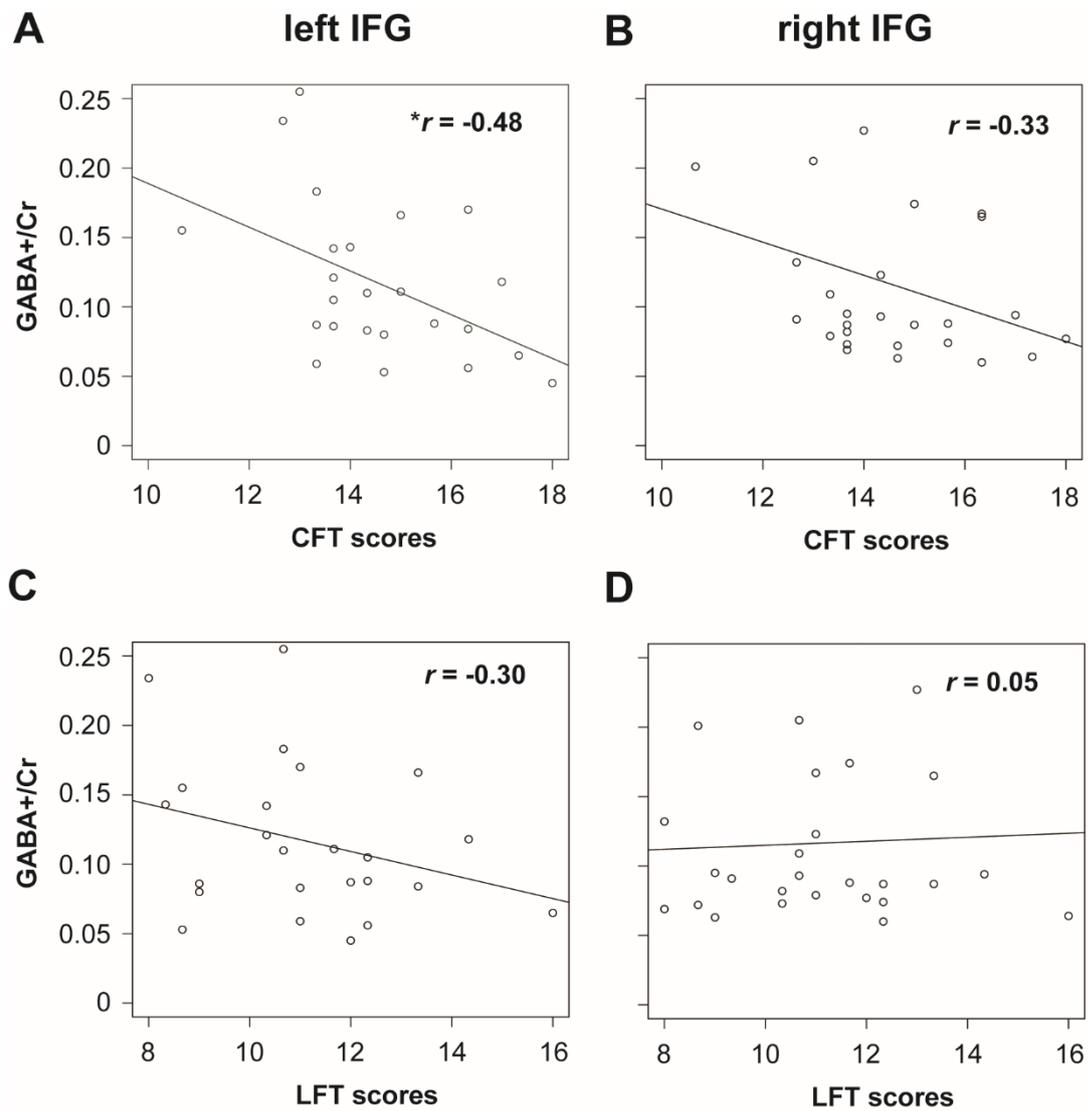
**Figure 3. Correlation with verbal fluency scores.** Correlations between CFT scores and GABA+/Cr values (A) and between LFT scores and GABA+/Cr values (B) for the left IFG. Correlations between CFT scores and GABA+/Cr values (C) and between LFT scores and GABA+/Cr values in the right IFG (D). *: $p < 0.05$.

I also analyzed the GABA+ concentrations using the water signal as a reference. As a result, I found significant negative correlation between CFT scores and GABA+/H_2O in the left IFG (r = -0.42, p = 0.041) but not in the right IFG (r = -0.36, p = 0.064). In both regions, I did not find significant correlation between LFT scores and GABA+/H_2O (left IFG; r = -0.32, p = 0.13; right IFG; r = 0.02, p = 0.92).

The significant correlation was found between LFT and CFT scores (r = 0.66, p < 0.001), while no significant correlation was found between the left IFG GABA+/Cr and right IFG GABA+/Cr (r = 0.30, p = 0.16). To directly compare the correlation profiles in the bilateral IFG, I performed Chi-square tests for the correlation coefficients after z-transformation. The result showed no significant difference between the left and right IFG for the CFT scores (CFT scores, Z = 0.61, p = 0.55; LFT scores, Z = 1.21, p = 0.23).

2.4. Discussion

In this experiment, I performed a H^1-MRS (MEGA-PRESS) to quantify the GABA concentration in the bilateral IFG and examined whether the individual variability of verbal fluency depends on GABA concentrations in these regions. I found a significant negative correlation between CFT scores and GABA+/Cr values only in the left IFG, though GABA+/Cr values were not different in the bilateral IFG. No correlation was found between LFT scores and GABA+/Cr values. I also confirmed that the current results were independent of references used (Cr and H_2O). These results provide new evidence for the neurotransmitter basis of language processing.

Among the two sub-components of the verbal fluency task, Costafreda et al. (2006) reported in their systematic meta-analysis study that the left IFG is involved in both CFT and LFT. According to the authors, CFTs are more related to sematic processing, while LFTs are related to phonological processing. Such functional associations may explain why the left IFG has been attributed to both phonological and semantic processing (Heim et al., 2003; Newman & Twieg, 2001). Even within the left IFG, Costafreda et al. further showed that the activation peaks in the LFT were located in more dorsal areas compared to those in the CFT (Costafreda et al., 2006). In the present study, I placed the target voxel at the dorsal part of the bilateral IFG (BA 44). Although it is possible that the correlation of GABA concentrations only with the CFT

is due to the voxel location, the significant correlation between the LFT and CFT scores suggests that the correlation of the GABA concentration with the LFT may also become significant with a larger sample size. The current MRS (MEGA-PRESS) measurement can be applied only to the single voxel each time. I should have performed another MRS measurement to a voxel including BA 45.

Recent MRS studies have reported negative correlations between various behavioral measures and GABA concentrations in corresponding cortical regions (Boy et al., 2010; Muthukumaraswamy et al., 2009; Puts et al., 2011; Terhune et al., 2014). However, there have been relatively few MRS studies aimed at higher cognitive functions such as language. In a recent study by Homan et al. (2014), patients with schizophrenia and normal control participants performed a verbal fluency task while MRS measurements were performed in the individual ROIs in the left IFG, which was activated during the task. Although the authors did not report the GABA concentrations, they found lower levels of N-acetyl-aspartate in the schizophrenia patients compared to the normal controls, supporting the idea of a neurotransmitter basis for linguistic function in this region. Michels et al. (2012) reported task-related temporal changes of GABA concentrations in the left dorsolateral prefrontal cortex during a Sternberg working memory (WM) task. However, because this WM task required participants to maintain five or seven letters in memory, the results may reflect the temporal modulation of the left frontal GABA concentration induced by verbal processing.

The leftward structural asymmetry of the IFG has been reported in previous studies (Amunts et al., 2003; Foundas et al., 1996). However, I found no difference between GABA+/Cr values in the left and right IFG. It seems that resting state GABA concentration is not a critical factor determining asymmetry in the IFG. In a previous study on post-mortem human brains, no left-right asymmetry of GABA concentration was found in the cortical regions, but was found in the substantia nigra (Rossor et al., 1980). A subsequent study suggested that individual variability of left-right asymmetry in the various neurotransmitters is rather important and characterizes each brain (Glick et al., 1982). I found a significant correlation between GABA+/Cr and CFT scores in the left IFG and not in the right IFG, but no significant difference of correlation coefficients was found. Although the present result suggests a contribution of GABA

concentration in the left IFG to linguistic performance, the asymmetric pattern of GABA concentration is not clarified.

The absence of GABA concentration difference between left and right IFG is also contrary to the idea of neurotransmitter asymmetry of the left and right IFG. If GABA concentration in the left IFG is correlated with CFT, and if GABA concentration in the left IFG equals to that of the right IFG, it is followed that the GABA concentration in the right IFG should also be correlated with the CFT. Although I did not find a significant correlation between GABA concentrations in the left and right IFG, it is not clear whether the GABA concentrations in these two regions are independent or not. Further examination is necessary to claim the neurotransmitter asymmetry in the IFG.

It is important to note a few limitations. Firstly, it is possible that the relatively small voxel size ($20 \times 20 \times 20$ mm^3) and small average number of channels used (64 in each region for both ON and OFF spectra) increased the variability of the spectra data, which might affect the data quality. Secondly, for all participants I first measured MRS data in the R. IFG, and the in the L. IFG. The lack of the counterbalancing of measurement bias might affect the hemispheric difference of the IFG. Lastly, although the correlation value of CFT scores in the left IFG ($P = 0.017$) did not reach statistical significance with multiple comparisons, the effect size ($r = -0.48$) was close to *large* according to Cohen's suggestion ($r = 0.1$, *small*; $r = 0.3$, *medium*; $r = 0.5$, *large*) (J. Cohen, 1988). Therefore, I consider that the correlation results in the present study reflect a meaningful property of the IFG. I conclude that GABA concentrations in the left IFG provide a neurotransmitter basis for linguistic performance.

Chapter 3. Cross-domain structural priming effect between linguistic and numerical tasks

（本チャプターは今後刊行予定であるため、除外する。）

Chapter 4. Cross-domain repetition suppression effect in the IFG

（本チャプターは今後刊行予定であるため、除外する。）

Chapter 5. The role of the left IFG and reward system for linguistic and numerical tasks

5.1. Introduction

The involvement of the left IFG for linguistic task has been repeatedly reported (Makuuchi et al., 2009; Musso et al., 2003; Pallier et al., 2011; Stromswold et al., 1996), while that for numerical task is controversial (Amalric & Dehaene, 2016; Ansari, 2016; Fedorenko et al., 2011; Maruyama et al., 2012; Monti et al., 2012; Nakai & Sakai, 2014). In the previous study of Nakai and Sakai (2014), I used numerical sequence stimuli ("1, 3, 5, …") and revealed that the numerical stimuli with complex syntactic structures induced activations in the left IFG. Hung et al. (2015), however, claimed that in the previous study of Nakai and Sakai (2014) the number of operators (addition and subtraction) also increased with syntactically complex numerical stimuli. To answer this issue, linguistic stimuli with similar structures as numerical sequence are necessary. Here I prepared Japanese (Hiragana) letter sequences as well as numerical sequences. Hiragana has a serial, alphabet-like arrangement, which is a perfect counterpart to numbers as fully familiarized sequential symbols for native Japanese participants.

The importance of functional integration between the left IFG and human reward system during learning period of language has been suggested (Ripollés et al., 2014). Recent neuroimaging studies revealed the involvement of the caudate, a part of the reward system, in processing of motivation (Delgado, Stenger, & Fiez, 2004; Murayama, Matsumoto, Izuma, & Matsumoto, 2010). The caudate has been also reportedly involved in reward anticipation and reward-based learning (Haruno et al., 2004; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006). Consistent with human neuroimaging studies, activations in primate caudate neurons were modulated by motivational context (Kawagoe, Takikawa, & Hikosaka, 1998). It is thus an important question whether the activations in the caudate reflect motivation for learning linguistic and numerical problems.

A theoretical model proposed that satisfaction induced by solving problems, or sense of accomplishment (SA), enhances motivation for learning (Keller, 1987). Another study indicated that the most obvious and ubiquitous source of intrinsic motivation for learning would be to offer challenging activities to individuals (Lepper & Hodell., 1989). It has

been proposed that tasks with intermediate difficulty, which are not too easy, nor impossibly hard, will induce the largest satisfaction for children after solving the problem (Harter, 1978a, 1978b). The efficacy of the optimal challenge for children's motivation has also been suggested in the field of physical activity (Weiss, 2000). It is likely that individuals do not feel SA under the direct instruction because the task becomes too easy for them, and that the minimal instruction is ineffective due to a lower successful rate. I hypothesized that even when solving the same problem participants would feel the greatest SA with indirect instruction, which may impose intermediate difficulty for them. Although behavioral studies have suggested the effect of challenging activities on SA (Harter, 1978a; Lepper & Hodell1989), no neuroimaging study has reported the neural basis of such sensation.

In the present experiment, I sought to examine the neural basis of SA. In particular, I focused on the bilateral caudate, and anterior cingulate cortex (ACC). I expect that the caudate is more related to the SA, because the SA may be relevant to the motivation-enhancing aspect in learning. Several studies have reported that the ACC is involved in processing emotional salience (Goldstein et al., 2009; Phan et al., 2004), and it is possible that emotional salience is tied up with the satisfaction of problem-solving as a secondary effect. I predicted that activations in the bilateral caudate would principally reflect SA induced by problem-solving.

To investigate the effects of instruction upon neural activations, I prepared three levels of instructions (No hint, Indirect hint, and Direct hint) for the same insight-problem types, and measured neural activations using fMRI (Figure 4, 5). I enhanced the instruction effect by using insight-problem solving task which is difficult to solve without any instructions. To ensure that the instruction effect is not domain-specific, I created insight-problems with number and letter notations. I prepared a basic control with smaller SA by including a time-counting task that did not require problem solving (simply called "Control"). I first performed a whole-brain analysis to reveal brain regions where activations are modulated by SA. According to the hypothesis, I then examined signal changes in the anatomically defined ROIs of the bilateral caudate and ACC. The current results indicate that SA is represented in the reward system, and that the Indirect instruction effectively induces such sensation.
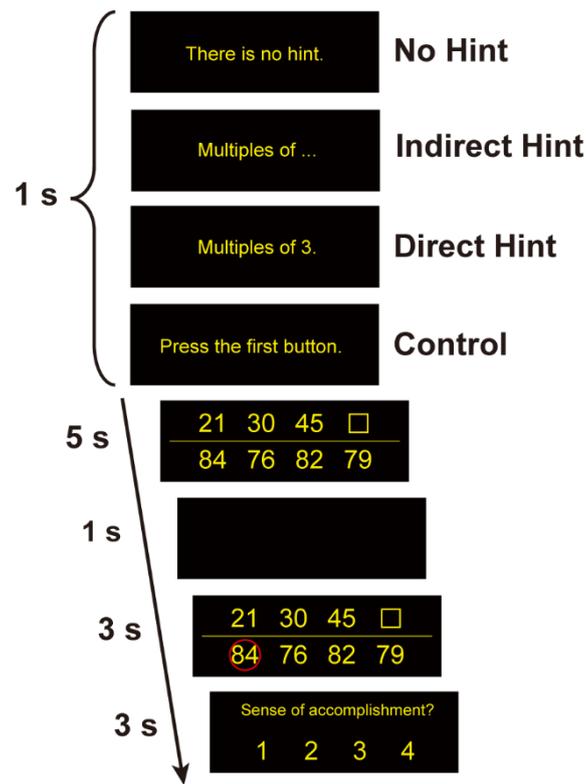
Figure 4. Experimental design. One of the four types of instructions was presented for 1 s, followed by the problem stimulus. Participants performed Number/Letter completion or button-press control tasks during the 5 s interval. After 1 s of a blank screen, 3 s of correct-answer feedback stimulus was presented. After the feedback stimuli, the participants rated their sense of accomplishment (SA) during the 3 s interval. Here I have shown an example of a Number notation task (translated in English).
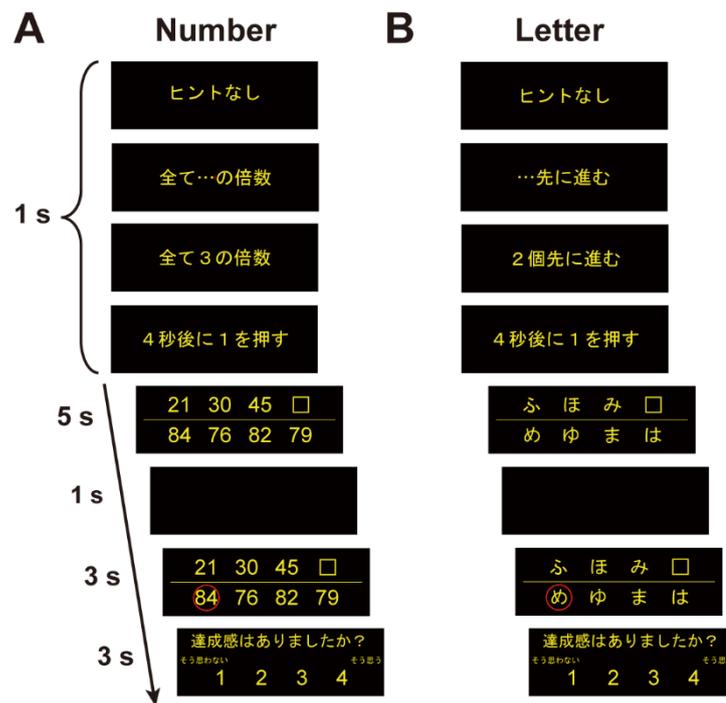
Figure 5. Examples of original stimuli. Original Japanese stimuli are shown in Number (A) and (B) Letter notations.

5.2. Methods
5.2.1. Participants
Twenty-one native Japanese speakers (14 males, aged 18–24 years) participated in the current experiment. All participants were healthy and right-handed (laterality quotient: 60–100), according to the Edinburgh inventory (Oldfield, 1971). Prior to their participation in the study, written informed consent was obtained from all participants. This experiment was approved by the Ethics Committee of the University of Tokyo, Komaba.


5.2.2. Stimuli and Tasks
The problem stimuli were composed of two arrays, where the upper array contained three symbols and a blank square, and the lower array contains four symbols. Sixty-four problem stimuli were created in total. To avoid memory-based interactions between the Direct condition and the other two instruction types, I used different problem stimuli sets in the Direct condition compared to the No hint and Indirect conditions. Since no memory-based interaction was expected, we used the same problem stimulus sets in the Direct and Control conditions. Although the problem stimuli were not common among three instruction levels, all problem stimuli were created based on the same set of rules. For the number notation, I used one-digit or two-digits of Arabic numbers. For the letter notation, Japanese letters (Hiragana) were used. Since Japanese participants learned serial, alphabet-like arrangements of those letters through their normal educational course, these letters were a perfect counterpart to numbers as fully familiarized sequential symbols for native Japanese participants.

I prepared four conditions (No hint, Indirect, Direct, and Control) for each of two symbol notations (Number and Letter; Figure 4, 5). Each direct hint was constructed to give the participants full information regarding the procedures or concept that composed the number/letter array, which enabled the participants to solve the corresponding problems without considering any hidden rule. Indirect hints were systematically constructed by omitting one word from the corresponding direct hint. I chose to omit words that had parametric information (e.g., a base number "3" of multiples) necessary for solving the tasks, such that the participants had to find that parameter from various candidates. I made eight direct hint stimuli and corresponding indirect hint stimuli for

each of two notations, and prepared four different problem stimuli depicted by each direct hint.

The participants were asked to perform a symbol (number or letter) completion task. In each trial, one of four condition instructions was given to the participants 1 s before the presentation of problem stimuli (e.g., No hint: "No Hint", Indirect: "multiples of …," Direct: "multiples of 3," and Control: "Push the first button"). The participants selected a symbol from four candidates in the lower array to replace the blank square in the upper array, and pressed one of the four corresponding buttons (two buttons for the right hand, two for the left) as soon as possible within 5 s. I instructed the participants to press any one of four buttons even if they could not find an answer during this period, because I wanted to measure the reaction time in all of the trials for the fMRI analysis. In such case I allowed participants to press buttons after 5 s. For the example in Figure 4, three numbers were presented in the upper row (21, 30, 45). Since all of those numbers were multiples of 3, participants had to find the corresponding number (i.e., 84) which was also a multiple of 3, from the four candidates in the lower row (84, 76, 82, 79). They pressed the first button in this case. After 5 s, a black screen was presented for 1 s, followed by a correct-answer feedback stimulus for 3 s, where the correct symbol in the lower array was marked with a red circle. Finally, the participants rated SA they felt for the task during the 3 s interval (1: "Strongly disagree", 2: "Disagree", 3: "Agree", 4: "Strongly agree"). The next trial began after 1 s.

The fMRI experiment was composed of four scanning sessions, where each session consisted of 32 trials (four trials under each condition in each notation). Four conditions were randomly arranged as an event-related design. The total number of trials was 128, while the total number of problem stimuli was 64 (i.e., the same problem stimuli were used twice). At the end of each session, the accuracy of each condition in that session was displayed. To reduce the practice effect among stimuli, trials were counter-balanced across participants, and the order of trials was pseudorandomized under the following constraint: trials with the same problem stimuli were always arranged in the distant sessions (i.e., 1st and 3rd sessions, or 2nd and 4th sessions), and all trials in the same instruction type within a session were based on the different rules. After the fMRI experiment and outside the scanner, the participants were asked to assess the difficulty of all problem stimuli used in the fMRI experiment (i.e., intrinsic problem difficulty), in

four-point Likert scale ranging from 1: "Very easy" to 4: "Very difficult." by looking each problem stimulus with its direct-hint. These questionnaires were composed of twelve questions with a seven-point Likert scale, ranging from 1: "Strongly disagree" to 7: "Strongly agree."

The participants wore earplugs in the scanner. Stimuli were presented on a liquid-crystal display monitor (resolution: 1920 × 1080), so that participants viewed them through a mirror. For fixation, a small red cross was always shown at the centre of the screen. Reaction time (RT) was measured from the onset of the problem stimuli. The stimulus presentation and collection of behavioural data (accuracy and RT) were controlled using the Presentation software (Neurobehavioral Systems, Albany, CA).

5.2.3. Behavioural Data Analyses

Correct/incorrect trials were evaluated by the response matching regardless of RTs, and trials with no response were treated as incorrect trials. In the present study, I evaluated the efficacy of instructions by overall results of problem-solving including both successful and unsuccessful cases, not only by the successful cases with limited numbers. Therefore, I included both correct and incorrect trials for the SA calculation. I also divided all trials (apart from those under the Control condition) into three difficulty levels (Hard, Medium, and Easy) according to the length of RTs (long, intermediate, and short, respectively) for each participant. 32 trials with the shortest RTs were categorized as Easy, 32 trials with the longest RTs were categorized as Hard, and remaining 32 trials were categorized as Medium, wherein the same trial is not always included in the same difficulty level for each participant. Trials with no response were included in the Hard level in this analysis. For the direct comparisons among the conditions, I performed the comparison of the No hint and Indirect conditions, as well as that of the Indirect and Direct conditions, because of the interest in the effect of indirect hint on the behavioral data. Therefore, I applied Bonferroni correction for the multiple comparisons at significance level of $\alpha = 0.025$.

5.2.4. MRI Data Acquisition

The functional imaging was conducted on a 3.0 T scanner (MAGNETOM Prisma; Siemens, Erlangen, Germany) with a 20-channel head coil. I scanned 35 interleaved axial

slices that were 3.2-mm thick with a 0.8 mm gap, parallel to the anterior and posterior commissure line, using a T2*-weighted gradient-echo echo-planar imaging (EPI) sequence [repetition time (TR) = 2.0 s, echo time (TE) = 30 ms, flip angle (FA) = 90°, field of view (FOV) = 192×192 mm^2, resolution = 3×3 mm^2]. In a single session, I obtained 242 volumes following four dummy images, which allowed for the rise of the MR signals. For anatomical reference, high-resolution T1-weighted images of the whole brain (176 sagittal slices, $1 \times 1 \times 1$ mm^3) were also acquired from all participants with a Magnetization Prepared Rapid Acquisition Gradient Echo sequence (MPRAGE, TR = 2000 ms, TE = 2.9 ms, FA = 9°, FOV =256×256 mm^2).

5.2.5. fMRI Data Analyses

I performed fMRI data analyses using SPM12 statistical parametric mapping software (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm/). The acquisition timing of each slice was corrected using the middle slice as a reference for the EPI data. I realigned the EPI data from multiple sessions to the mean image across all sessions. Each participant's T1-weighted structural image was coregistered to the mean functional image generated during realignment, and then spatially normalized to the Montreal Neurological Institute (MNI) space with the new unified normalization-segmentation tool in SPM12. After spatial normalization, the resultant deformation field was applied to the realigned functional imaging data, resampled into 2 mm isotropic voxels. All normalized functional images were then smoothed using an isotropic Gaussian kernel of 8 mm full-width at half maximum. Low-frequency noise was removed by high-pass filtering at 1 / 128 Hz.

To separate the effect of reward anticipation and outcomes, I examined neural activations in both problem-solving and answer-feedback periods, by performing parametric modulation analyses of the fMRI data with two general linear models (GLM). Based on the previous studies of insight-problem solving (Jung-Beeman et al., 2004; Zhao et al., 2013), in the first GLM I used regressors composed of 3 s box-car functions, starting from 2 s before the button response for each trial (problem-solving period), convolved with a hemodynamic function. I selected this period to include the entire processes regarding SA and emotion related to the insight-problem solving. In the second GLM I used regressors composed of 3 s box-car function beginning with the onset of the correct-

answer feedback stimulus (answer-feedback period). In both of the above GLMs, I merged all four conditions and two notations, and included SA as a parametric modulator for each participant. Since fMRI data were analysed based on the participants' response, I excluded trials with no response from fMRI data analyses. The contrast image of positive modulation by the SA was obtained for each participants, and was used for intersubject comparisons in the second-level analysis. The statistical threshold was set to $p < 0.001$ for the voxel level, with $p < 0.05$ for the cluster level [topological False-discovery rate (FDR) correction for multiple comparisons] across the whole brain (Chumbley & Friston, 2009). I also reported some contrasts with uncorrected $p < 0.001$ which did not show significant activation with topological FDR correction, but might be suggestive for the interpretation of the present data.

For the region of interest (ROI)-based beta estimate analysis, I made two additional GLMs with four conditions for both Number and Letter notations, in both problem-solving and answer-feedback periods in the same way as above, but without SA regressors. Based on a priori interest in the regions related to the reward system and emotional saliency, I defined three anatomical ROIs for the bilateral caudate and ACC with the Automatic Anatomical Labeling atlas (Tzourio-Mazoyer et al., 2002), and three functionally defined ROIs of the bilateral caudate and ACC from the contrast of positive modulation by the SA. I then extracted the beta estimates averaged in each of those ROIs using the MarsBaR-toolbox (http://marsbar.sourceforge.net/). I also constructed an additional GLM with three difficulty levels (Hard, Medium, and Easy). To this end, I concatenated the scans from the separate sessions, and divided all trials (apart from those under the Control condition) into three difficulty levels according to the length of RT (long, intermediate, and short, respectively) for each participant. The effects of transition between sessions were taken into account with regressors of sessions. The direct comparisons among the conditions were performed in the same way as behavioral data, with Bonferroni correction for the multiple comparisons at significance level of $\alpha = 0.025$.

5.3. Results
5.3.1. Behavioural results
Regarding accuracy, two-way repeated measures analysis of variance (rANOVA) showed that the main effect of notation (Number and Letter) [$F(1, 20) = 15$, $p < 0.001$], and the

main effect of condition (No hint, Indirect, Direct, and Control) $[F(3, 60) = 111, p < 0.001]$, was significant (Figure 6A). The interaction was also significant $[F(3, 60) = 2.8, p = 0.048]$. Next I performed post-hoc paired t-tests among the four conditions on the notation-concatenated data. As a result, I found significant difference of accuracy between No hint and Indirect conditions ($p < 0.001$), between Indirect and Direct conditions ($p < 0.001$), and between the Direct and Control ($p = 0.0016$) (in summary, Control > Direct > Indirect > No hint). Similar patterns were observed when I analyzed the accuracy in Number and Letter notations separately (Figure 7A, B). As for RTs, the main effect of notation $[F(1, 20) = 11.1, p = 0.0034]$, and the main effect of condition $[F(3, 60) = 44.3, p < 0.001]$, was significant (Figure 6B). The interaction was also significant $[F(3, 60) = 5.8, p = 0.0015]$. Post-hoc paired t-tests showed significant difference of RTs between the No hint and Indirect conditions ($p < 0.001$), as well as between the Indirect and Direct conditions ($p < 0.001$) (in summary, No hint > Indirect > Direct). The direct comparisons between the Control and other conditions were not performed because in the Control condition participants were asked to always respond after 4 s of the problem onset. Similar patterns were observed when I analyzed the RTs in Number and Letter notations separately (Figure 7C, D). The number of trials where RT exceeded 5 s was 9.1 ± 5.1 (average ± s.d.) for the No hint condition, 5.7 ± 4.3 for the Indirect condition, and 2.2 ± 2.3 for the Direct condition (the total number of trials was 32 for each condition). Regarding SA, the main effect of notation $[F(1, 20) = 11.5, p = 0.0029]$, as well as the main effect of condition $[F(3, 60) = 14.7, p < 0.001]$, was significant (Figure 6C, D). The interaction was not significant $[F(3, 60) = 1.7, p = 0.18]$. To examine the influence of different instruction types on SA, I performed paired t-tests among the four conditions on the notation-concatenated data. As a result, I observed that the Indirect condition induced the largest SA ($p < 0.05$). Analysis in each notation revealed the significant difference of SA between Indirect and No hint conditions in the Letter notation (Figure 7E, F). Next I divided all conditions into correct and incorrect trials (Figure 6D, 20). By comparing SA in correct and incorrect trials, I found that those in the correct trials were significantly higher than those in the incorrect trials $[t(21) = 9.0, p < 0.001]$. I also examined SA in correct and incorrect trials under each condition (Figure 8). The one-way rANOVA in the correct trials showed the significant main effect condition $[F(3,60) = 30.87, p < 0.001]$. The post-hoc t-tests revealed that No hint and

Indirect conditions provided significantly larger SA than Direct condition ($p < 0.01$, Bonferroni corrected). Analysis within the incorrect trials was not performed because of missing values
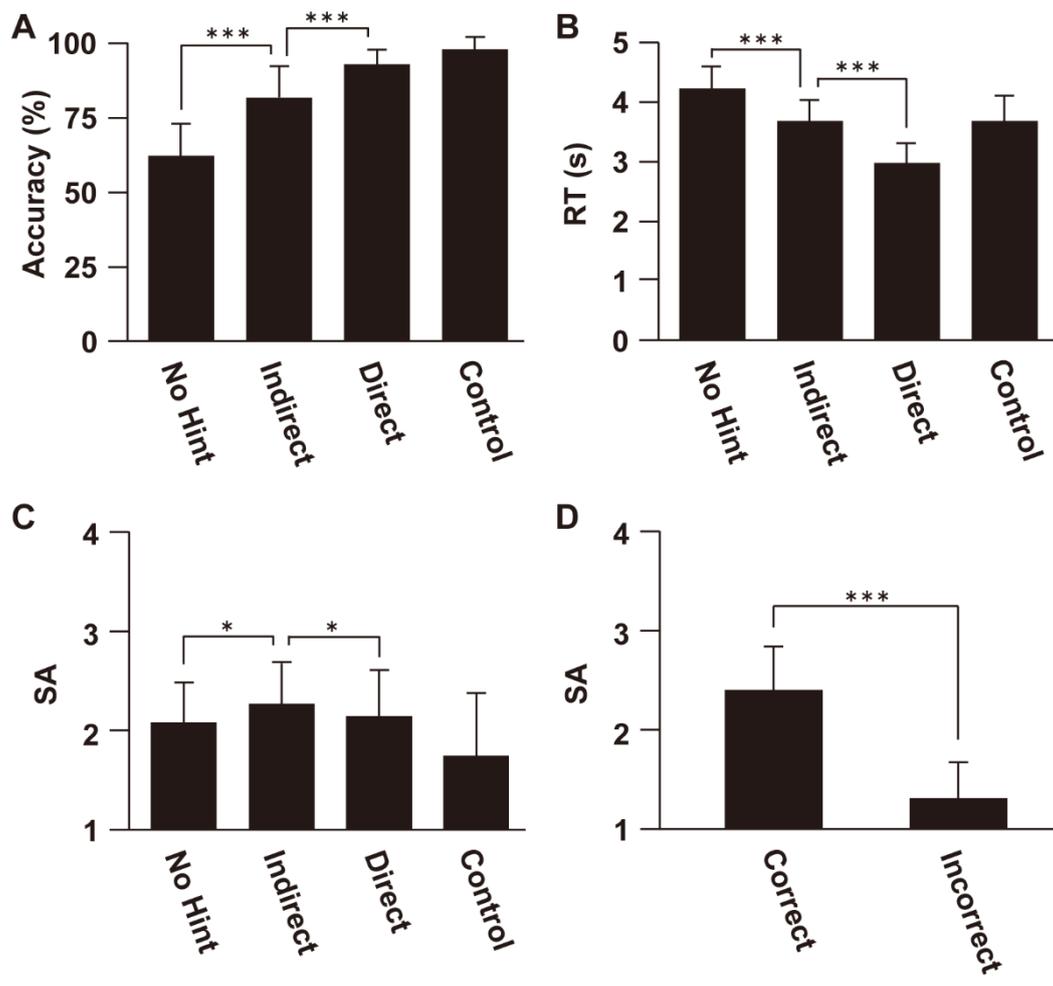
Figure 6. Behavioural data for the four conditions. Accuracy (A), RT (B), and SA (C) are shown for the four conditions (two notations concatenated). (D) SA in the correct and incorrect trials, averaged across No hint, Indirect, and Direct conditions. $*p < 0.05$, $***p < 0.001$ (Bonferroni corrected). Error bars, SD.
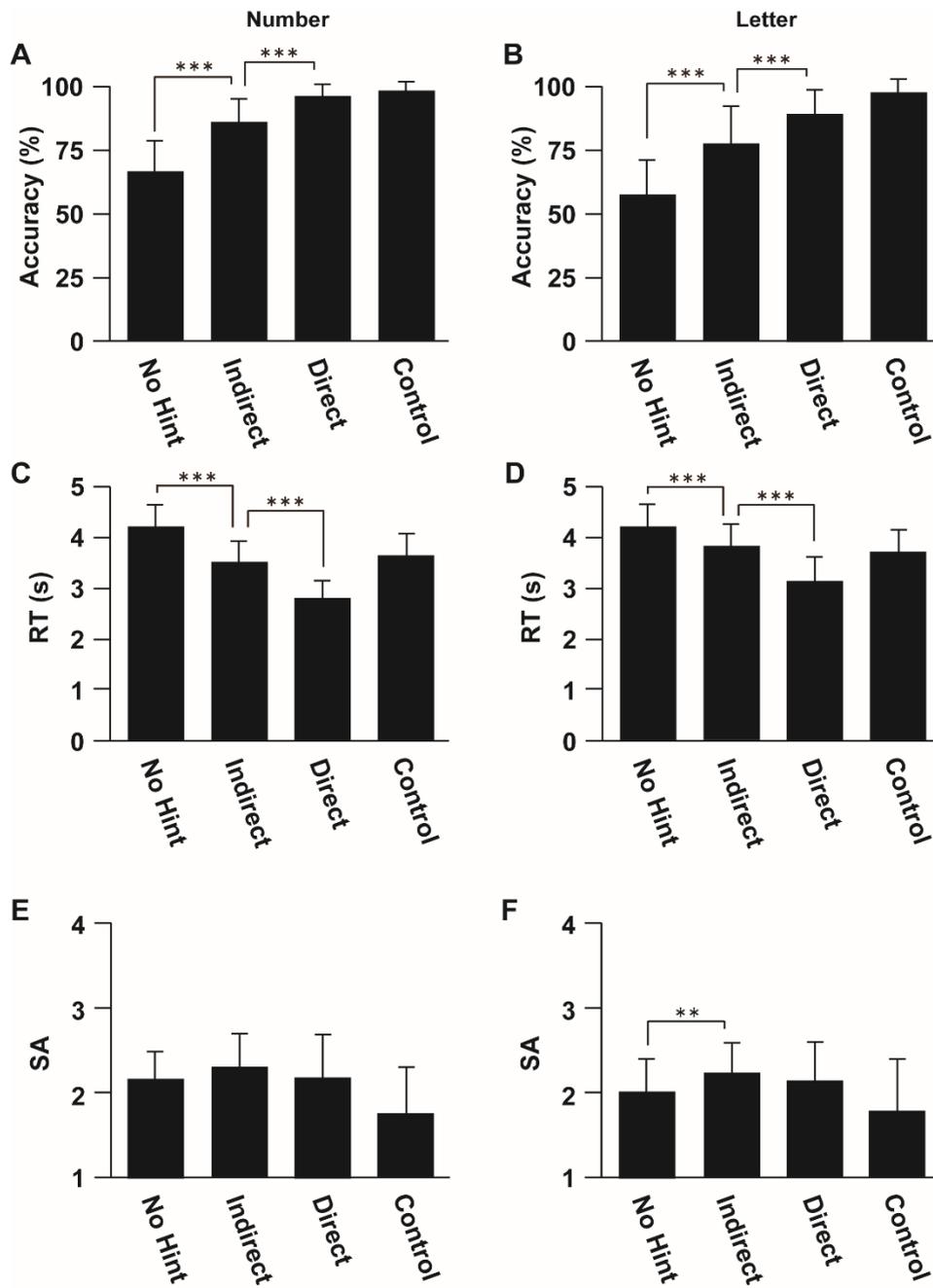
Figure 7. Accuracy, RT, and SA for both notations. Accuracy (A) (B), RT (C) (D), and SA (E) (F) are shown for four conditions, in both Number and Letter notations. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ (Bonferroni corrected). Error bars, SD.
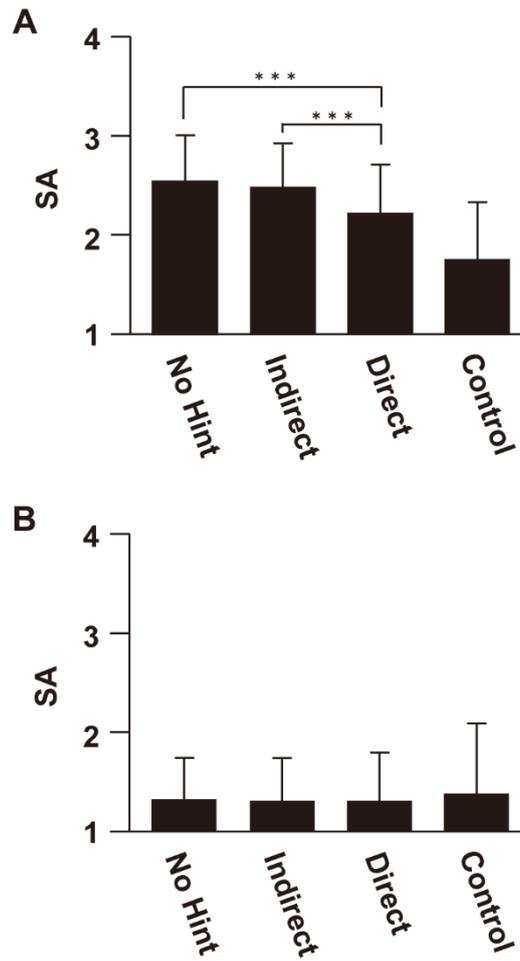
Figure 8. SA in the correct and incorrect trials. SA in the correct (A) and incorrect (B) trials are shown for the four conditions (two notations concatenated). ***$p < 0.001$ (Bonferroni corrected). Error bars, SD.

I further analysed accuracy and RT as a function of intrinsic problem difficulty obtained outside the MR scanner (Figure 9A, B, Figure 10A-D), and found that the three instruction types modulated both accuracy and RT. For the RT, two-way rANOVA showed a significant main effect of intrinsic problem difficulty [abbreviated as Difficulty, $F(3, 54) = 63.5, p < 0.001$] and condition [$F(2, 36) = 150.2, p < 0.001$], however the interaction was not significant [$F(6, 105) = 1.8, p = 0.12$]. Regarding Accuracy, the main effects of intrinsic problem difficulty [$F(3, 54) = 28.4, p < 0.001$] and condition [$F(2, 36) = 67.3, p < 0.001$] were significant, however the interaction was not significant [$F(6, 105) = 1.5, p = 0.12$]. The overall intrinsic problem difficulty was different between two notations [$t(21) = 3.6, p < 0.001$, Figure 10E].

To diminish the influence of the intrinsic problem difficulty and individual variability, I divided all trials (apart from those under the Control condition) into three difficulty levels (Hard, Medium, and Easy) according to the length of RTs. I again performed paired t-tests among three difficulty levels, and found that the Medium level (with intermediate difficulty) had the largest SA ($p < 0.001$, Bonferroni corrected, Figure 9C). Within the Number notation, I again found that the Medium level had the largest SA ($p < 0.01$, Bonferroni corrected, Figure 11A). Within the Letter notation, only the difference between the Hard and Medium levels was significant ($p < 0.001$, Bonferroni corrected, Figure 11B).
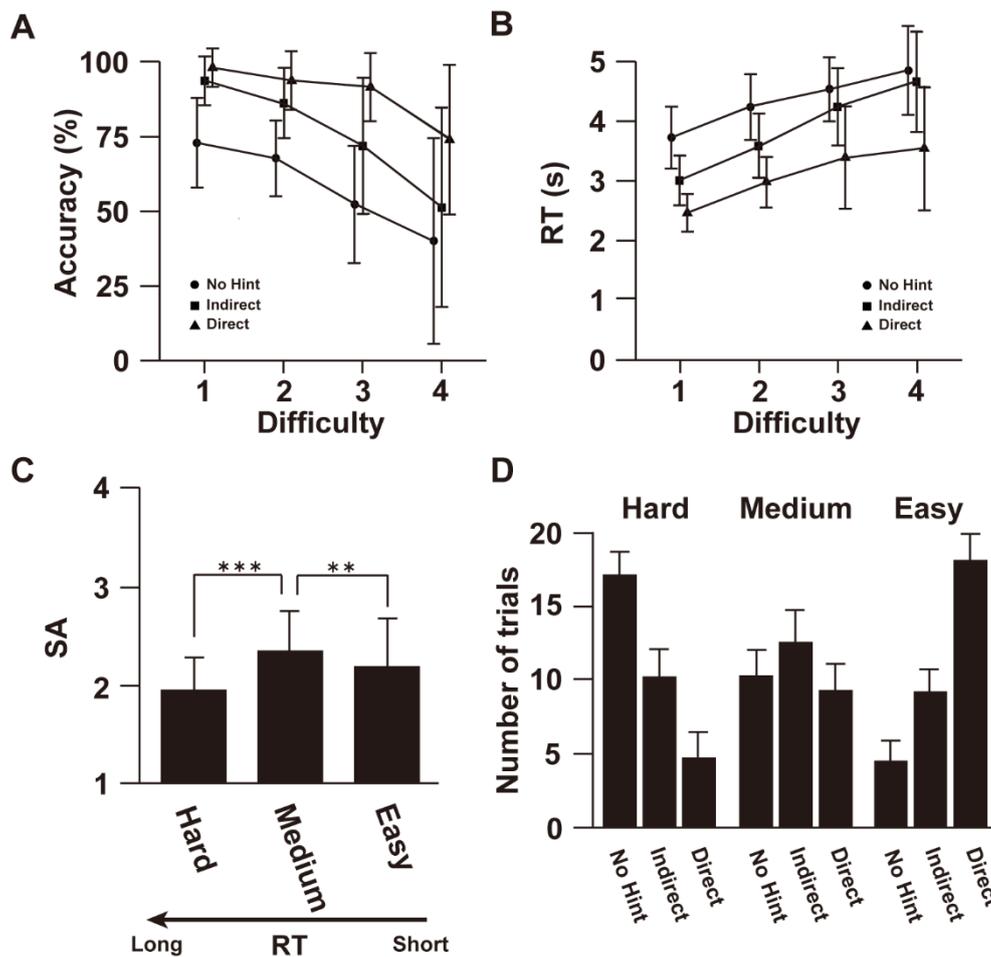
Figure 9. Effect of intrinsic problem difficulty. Accuracy (A) and RT (B) for three instruction levels are shown as a function of intrinsic problem difficulty (Difficulty) as measured outside the MR scanner. (C) SA for the three difficulty levels, divided according to the RT differences for each participant. (D) The average number of trials for each of three instructions included in each difficulty level. ***$p < 0.001$, **$p < 0.01$ (Bonferroni corrected). Error bars, SD.
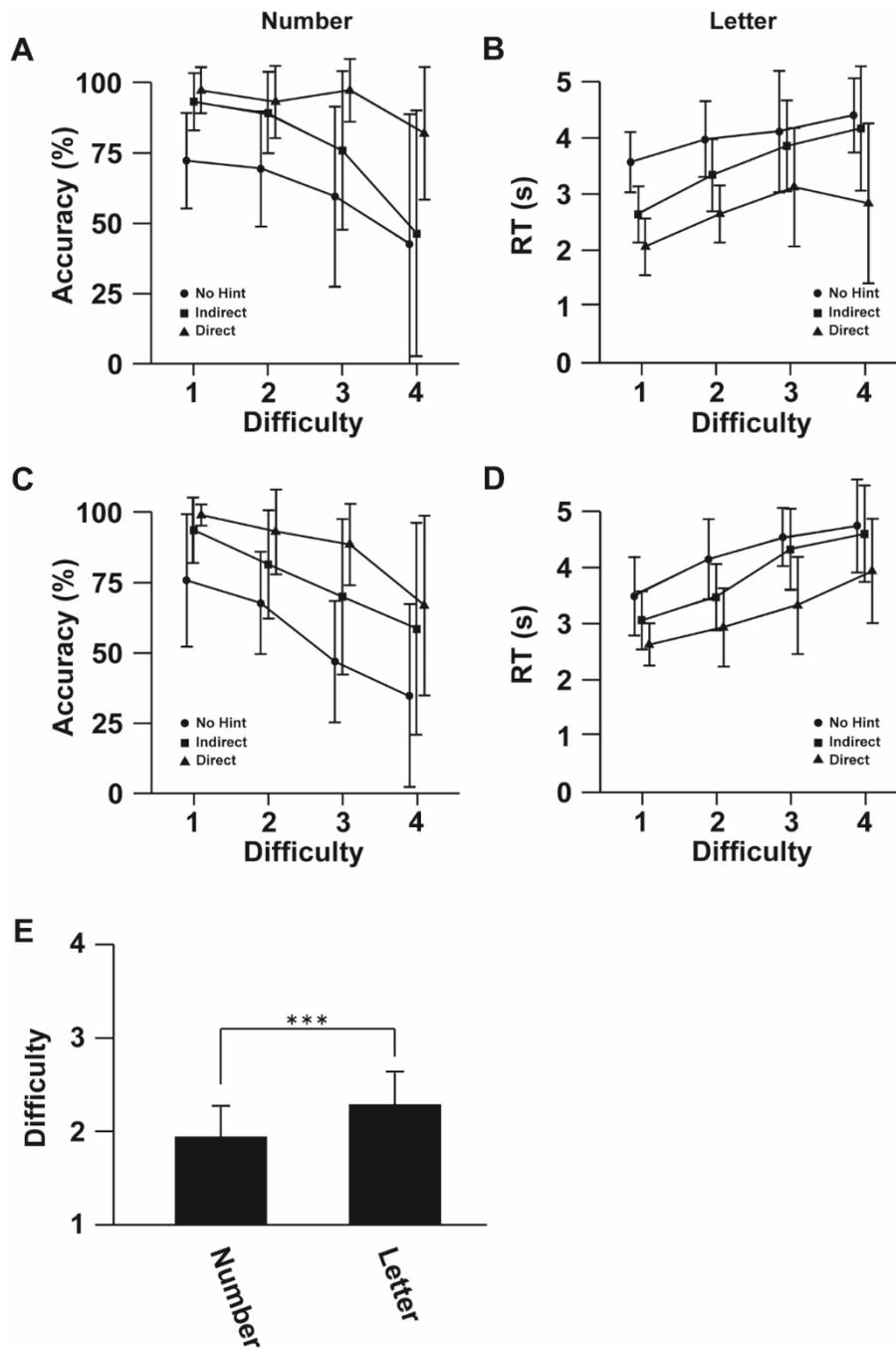
Figure 4. Effect of intrinsic difficulty for both notations. Accuracy (A) (C) and RT (B) (D) in three levels of instructions are shown as a function of intrinsic problem difficulty (Difficulty), analysed separately for Number and Letter notations. (E) Difference of Difficulty (averaged for all conditions) between two notations. ***$p < 0.001$. Error bars, SD.
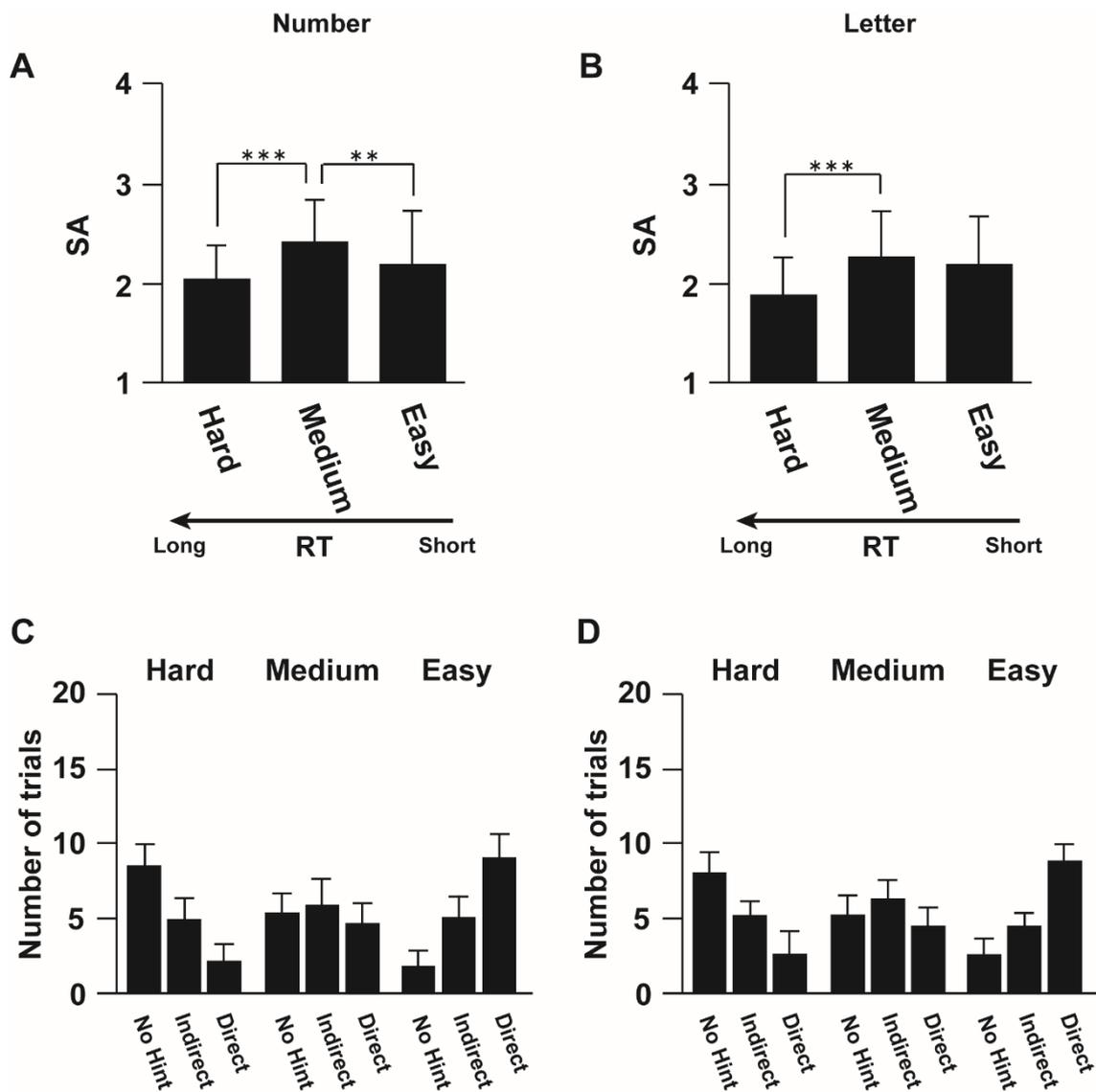
Figure 11. Three difficulty level analysis shown for both notations. (A) (B) SA for the three difficulty levels, divided according to the RT difference, and (C) (D) the average number of trials for three instructions included in each difficulty level, analyzed separately for Number and Letter notations. **$p < 0.01$, ***$p < 0.001$ (Bonferroni corrected). Error bars, SD.

5.3.2. fMRI results

In the following fMRI analysis, I concatenated Number and Letter notations. To confirm that predicted signals of problem-solving period and answer-feedback period were well-separated, I performed a correlation analysis of regressors. For each participant, I convolved a hemodynamic response function with box-car functions in the problem-solving period and feedback period, and calculated correlation coefficient between these two convolved functions. The average correlation coefficient was -0.20 ± 0.09 (all conditions combined). This low coefficient value suggests that signal overflow effect was limited.

In order to know the brain regions generally related to problem-solving activity in the experiment, I examined neural activation without including SA as regressors, and compared activation under each instruction (Figure 12, Table 1). I found significant activation in the bilateral IPL, occipital cortex in the No Hint – Control contrast. In the Indirect – Control contrast, I additionally found activation in the left IFG and left lateral premotor cortex (LPMC). Although no significant activation of the left IFG was found in the Direct – Control and No hint – Control contrast, direct comparisons of signal changes revealed significantly larger activations in the Direct and No hint conditions than Control condition ($p < 0.001$, Bonferroni corrected).
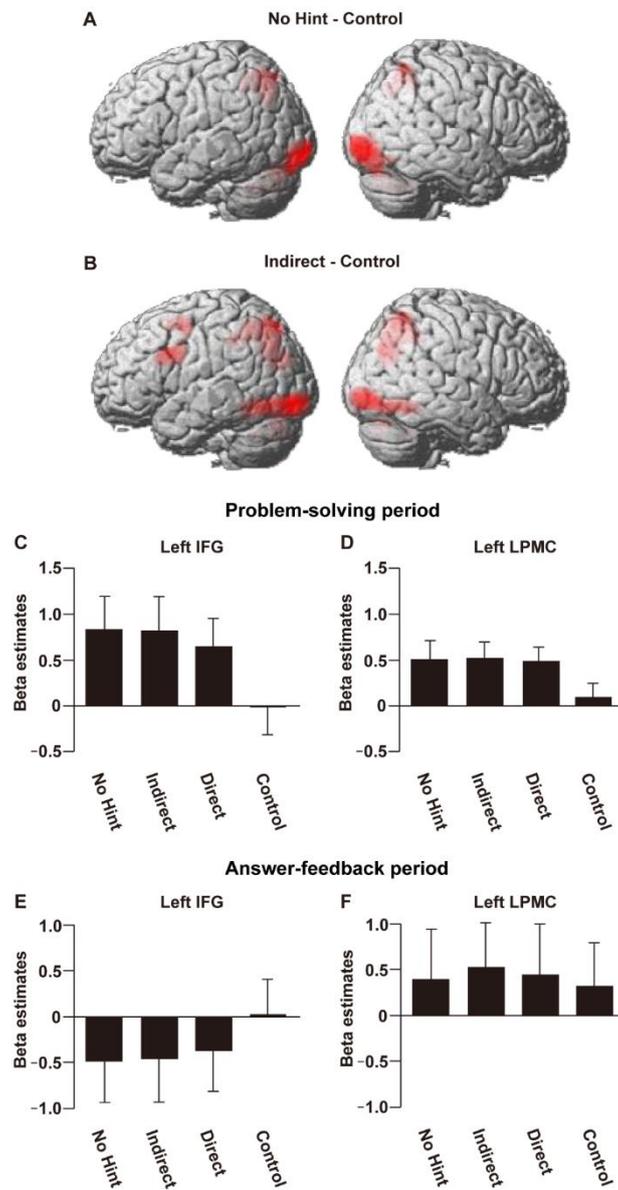
Figure 12. Neural activations without a parametric modulation regressor. The cortical activation maps of No hint – Control (A) and Indirect – Control (B) were projected onto the standard brain ($p < 0.001$ for voxel level, $p < 0.05$ for cluster level, with topological FDR correction). See Table 1 for stereotactic coordinates. Beta estimates for four conditions were extracted from functionally defined ROIs (left IFG and LPMC) of Indirect – Control contrast (C-F). Error bars, SD.

Brain region	BA	Side	x	y	z	Z-value	Voxels
No hint – Control							
IPL	7	L	–24	–62	48	5.3	883
	7	R	28	–58	52	4.2	452
			26	–68	44	4.0	*
MOG	18	L	–20	–94	–8	> 8.0	4109
		R	26	–90	–6	> 8.0	*
OP	17	M	–2	–74	–24	6.0	*
Indirect – Control							
LPMC	6	L	–28	–2	54	5.8	446
IFG	44	L	–42	4	30	5.1	702
			–54	12	34	3.3	*
	45	L	–38	14	26	4.9	*
IPL	7	L	–24	–62	48	5.8	1846
			–16	–70	54	4.9	*
	7	R	28	–58	50	4.5	1348
	40	R	26	–46	42	4.4	*
LOG	19	L	–26	–68	28	4.9	*
	19	R	32	–72	34	5.1	*
MOG	18	R	28	–90	–6	5.7	1248
OTG	37	R	44	–62	–12	4.8	*
MOG	18	L	–24	–92	–10	5.6	1644
OTG	37	L	–46	–60	–10	5.4	*
LOG	19	L	–38	–80	–10	5.2	*
Cerebellum		M	–4	–74	–26	5.1	843
			–2	–56	–36	4.6	*

Table 1. Direct comparison among instruction levels. Stereotactic coordinates (*x, y, z*) in the MNI space (mm) are shown for each activation peak of *Z*-values. IPL, inferior parietal lobule; OP, occipital pole; IFG, inferior frontal gyrus; LPMC, lateral premotor cortex; OTG, occipitotemporal gyrus. The region with an asterisk is included within the same cluster shown one row above.

To identify the cortical regions reflecting notation-independent SA during the problem-solving period, I performed a one-sample t-test with images of positive modulation by SA for each participant (Figure 13A, B, Table 2). I found significant modulation including the bilateral caudate, ACC [BA 11], and posterior cingulate cortex (PCC, BA 23). I also found modulation in the left ventrolateral prefrontal cortex (vlPFC, BA 47), vmPFC (BA 11), right middle occipital gyrus (MOG, BA 18), and right lateral occipital gyrus (LOG, BA 19). No region showed significant negative modulation with SA. When I performed the whole brain analysis in each instruction level (uncorrected $p < 0.001$, Figure 14), I found significant positive modulation with SA in the ACC, PCC, and ventral putamen under the No Hint condition, ACC under the Indirect condition, and bilateral caudate under the Direct condition. None of those regions were modulated under the Control condition.
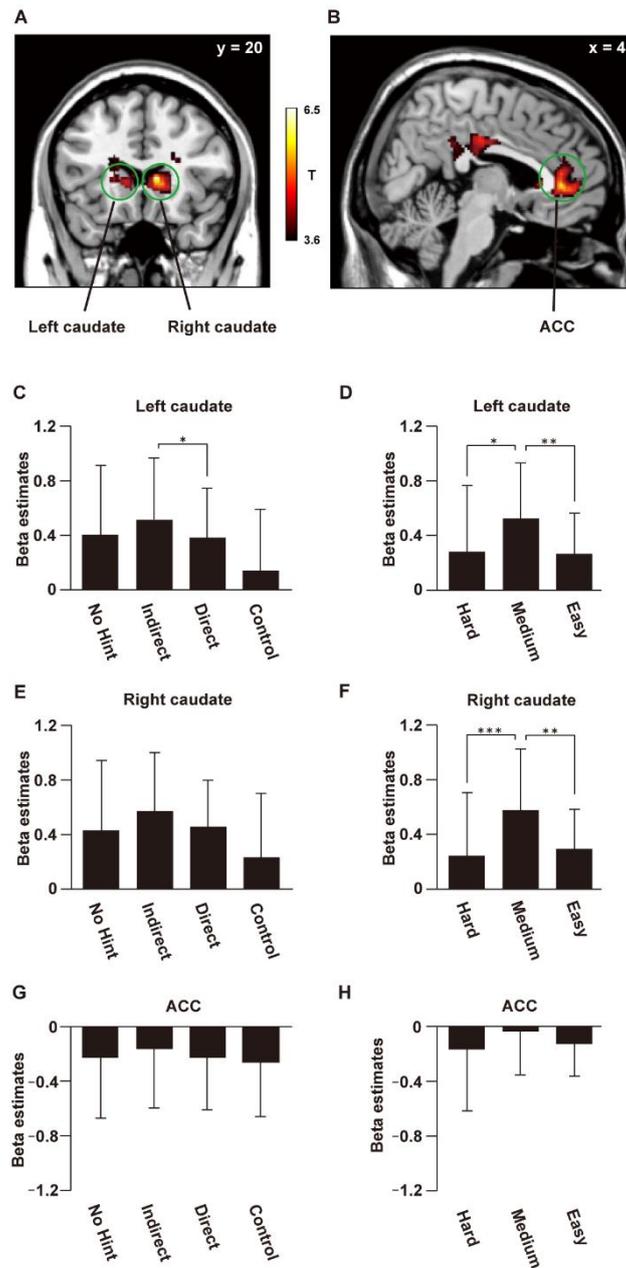
Figure 13. Neural activations parametrically modulated by SA. The cortical activation map was projected onto coronal (A) and sagittal (B) plains ($p < 0.001$ for voxel level, $p < 0.05$ for cluster level, with topological FDR correction). See Table 2 for stereotactic coordinates. Beta estimates for four conditions were extracted from anatomical ROIs of the left caudate (C), the right caudate (E), and ACC (G). In the same ROIs, I extracted beta estimates for the three difficulty levels divided according to the length of RT. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ (Bonferroni corrected). Error bars, SD.

Brain region	BA	Side	x	y	z	Z-value	Voxels
vmPFC	11	L	–18	42	–6	4.2	503
			–20	50	–2	3.8	*
vlPFC	47	L	–34	44	–6	4.1	*
ACC	11	M	4	38	2	4.4	1854
Caudate		L	–20	24	4	4.7	*
		R	10	20	6	4.6	*
		R	18	–8	22	3.8	414
			26	–24	24	3.7	*
Thalamus		R	18	–16	16	3.6	*
PCC	23	M	–2	–20	34	4.7	622
			–6	–6	28	4.1	*
MOG	18	R	16	–98	0	3.9	867
			22	–82	–6	3.7	*
LOG	19	R	36	–86	–6	3.6	*

Table 2. Regions where activations were modulated by SA. Stereotactic coordinates (*x, y, z*) in the MNI space (mm) are shown for each activation peak of Z-values. ACC, anterior cingulate cortex; PCC, posterior cingulate cortex; vmPFC, ventromedial prefrontal cortex; vlPFC, ventrolateral prefrontal cortex; MOG, middle occipital gyrus; LOG, lateral occipital gyrus; BA, Brodmann's area; L, left hemisphere; R, right hemisphere; M, medial. The region with an asterisk is included within the same cluster shown one row above.
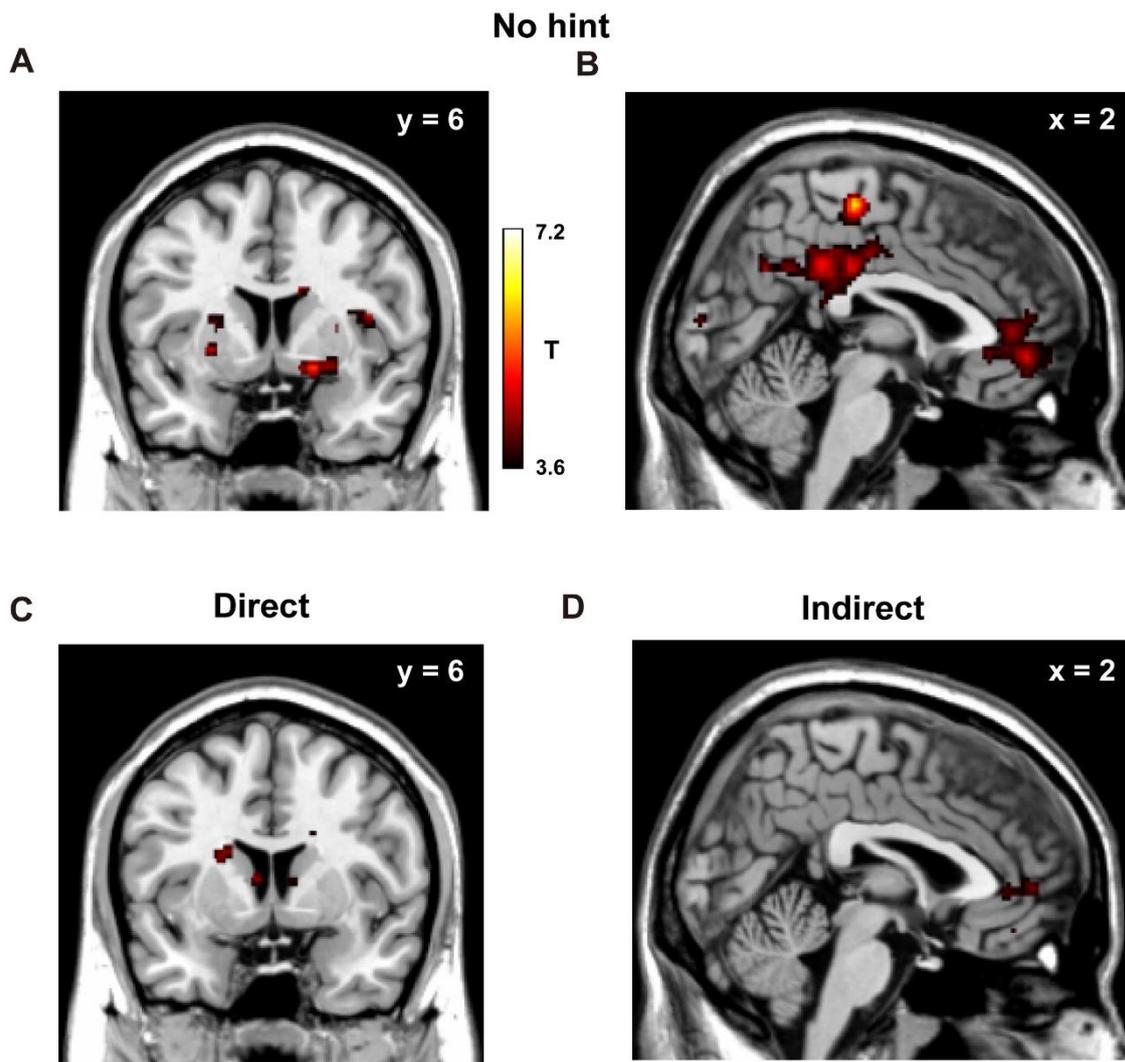
**Figure 14. Positive modulation with SA in each instruction level.** The cortical activation map was projected onto the coronal (A) and sagittal (B) plains for the No hint condition, coronal plain for the Direct condition (C), and sagittal plain for the Indirect condition (D) (uncorrected $p < 0.001$).

According to a priori interest in the reward and emotion-related systems, next I examined beta estimates in the anatomical ROIs of the bilateral caudate and ACC (Figure 13C, E, G, 12). To confirm that the activation pattern is independent of notations, I used two-way rANOVA. As a result, I found a significant main effect of condition in the bilateral caudate [left caudate: $F(3, 60) = 6.2$, $p < 0.001$, right caudate: $F(3, 60) = 5.0$, $p = 0.0037$]. In these regions, the main effect of notation [left caudate: $F(1, 20) = 1.9$, $p = 0.18$, right caudate: $F(1, 20) = 1.2$, $p = 0.29$], as well as the interaction [left caudate: $F(3, 60) = 0.77$, $p = 0.52$, right caudate: $F(3, 60) = 0.42$, $p = 0.73$], was not significant. In the ACC, I did not find the significant main effect of condition [$F(3,60) = 0.34$, $p = 80$], the main effect of notation [$F(3,60) = 0.0080$, $p = 0.93$], or the interaction [$F(3,60) = 0.55$, $p = 0.65$]. Paired t-tests for the notation-concatenated data revealed a significant difference between the Indirect and Direct conditions in the left caudate ($p = 0.0070$).

As above, I divided all trials (apart from those under the Control condition) into three difficulty levels (Hard, Medium, and Easy) according to the length of RTs (Figure 13D, F, H), since I confirmed through the behavioural data that this division should be sensitive to the individual variability and effect of intrinsic problem difficulty. This analysis revealed that Medium level trials showed significantly larger activations than other levels in the bilateral caudate ($p < 0.01$), consistent with the behavioural data. Activations in the ACC were negative (Figure 13G-H, 15E-F), and I did not find a significant main effect of condition [$F(3, 60) = 0.34$, $p = 0.80$], or the main effect of notation [$F(1, 20) = 0.008$, $p = 0.93$]. The interaction was not significant [$F(3, 60) = 0.55$, $p = 0.65$]. Even after dividing all trials into three difficulty levels, no significant difference in activation was found among those levels in the ACC [$F(2, 40) = 1.7$, $p = 0.20$, Figure 13H). The difference between the caudate and ACC was further indicated by analysis of the answer-feedback period, which showed activations modulated by SA in the ACC, but not in the caudate (uncorrected $p < 0.001$, Figure 16, Table 3).
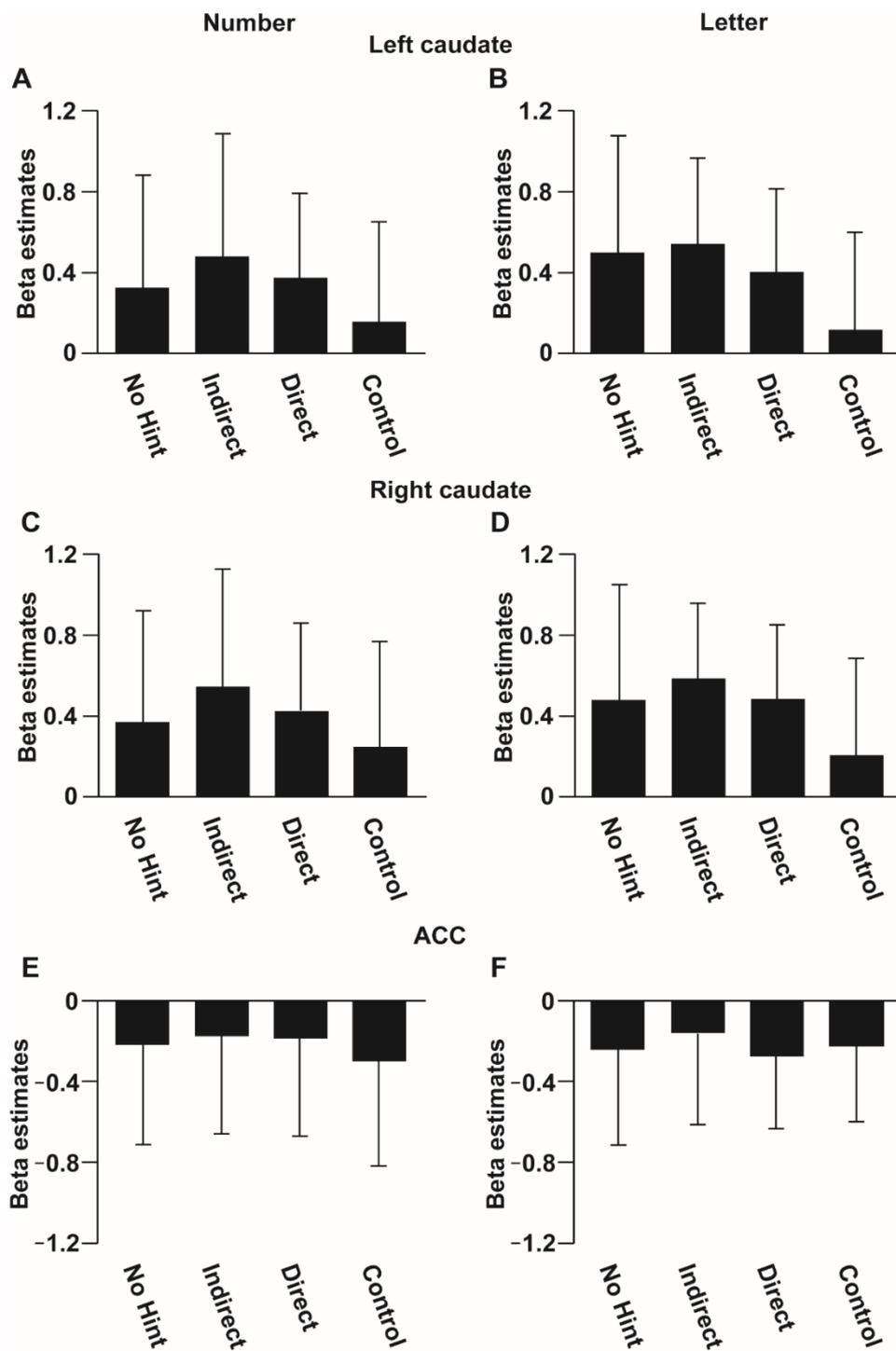
Figure 15. Beta estimates for both notations. Beta estimates were extracted from the anatomical ROIs of the left caudate (A, B), right caudate (C, D), and ACC (E, F), analysed separately for Number and Letter notations. Error bars, SD.
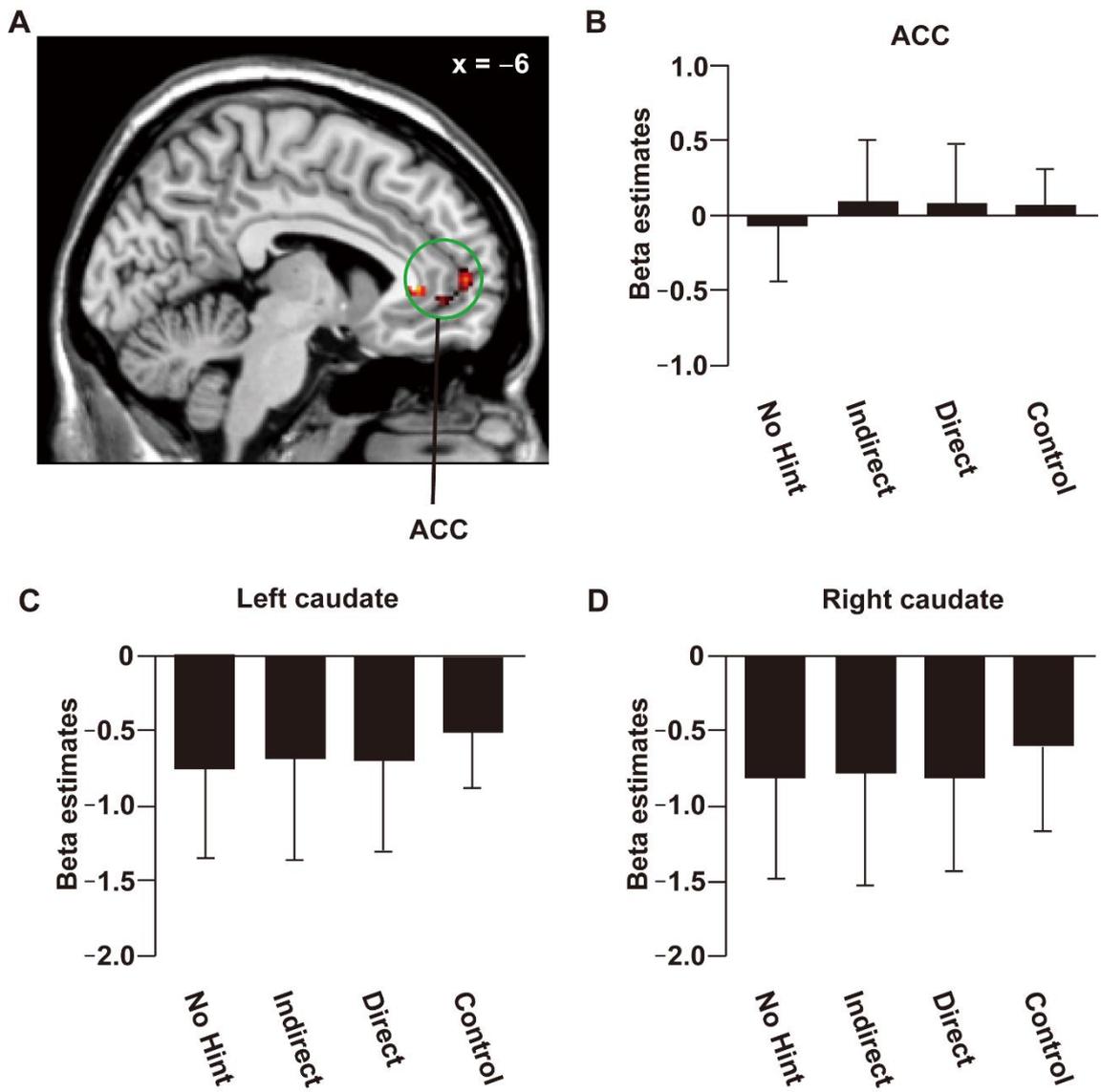
Figure 16. Activation modulated by SA in the answer-feedback period. The cortical activation map was projected onto the sagittal plain (uncorrected $p < 0.001$). See Table 3 for the stereotactic coordinates.

Brain region	BA	Side	x	y	z	Z-Value	Voxels
vmPFC	11	R	16	50	–4	3.1	1
dlPFC	9	L	–12	56	34	3.5	23
			–14	46	48	3.3	5
ACC	11	M	–6	34	2	3.8	125
			–6	56	6	3.5	*
			–2	46	–2	3.4	*
PCG	4	L	–34	–24	54	3.5	50
OP	17	R	14	–88	0	3.2	2
LOG	19	R	14	–54	–18	3.5	14

Table 3. Activations modulated by SA in the answer-feedback period. Stereotactic coordinates (*x, y, z*) in the MNI space (mm) are shown for each activation peak of *Z*-values. dlPFC, dorsolateral prefrontal cortex; PCG, postcentral gyrus; OP, occipital pole.

To show that the results of SA modulation is robust regarding the trial selections and effect of accuracy, I performed an additional parametric modulation analysis only for the correct trials (Figure 17). I found significant modulation in the right caudate, but not in the ACC (uncorrected $p < 0.001$). Next I analysed activations in the correct and incorrect trials separately (without the Control condition). In the problem-solving period, in each of the three regions I found that correct trials induced larger activation than incorrect trials [left caudate: $t(21) = 2.5$, $p = 0.011$, right caudate: $t(21) = 2.9$, $p = 0.004$, ACC: $t(21) = 3.4$, $p = 0.001$, Figure 18A-C]. In the answer-feedback period, in each of the three regions I again found that correct trials induced larger activation than incorrect trials [left caudate: $t(21) = 3.2$, $p = 0.002$, right caudate: $t(21) = 2.7$, $p = 0.007$, ACC: $t(21) = 3.5$, $p < 0.001$, Figure 18D-F]. Those results may indicate that activations in the ACC basically reflect the difference between correct and incorrect trials.

A y = 12

B x = 4

5.5

T

3.6

Right caudate
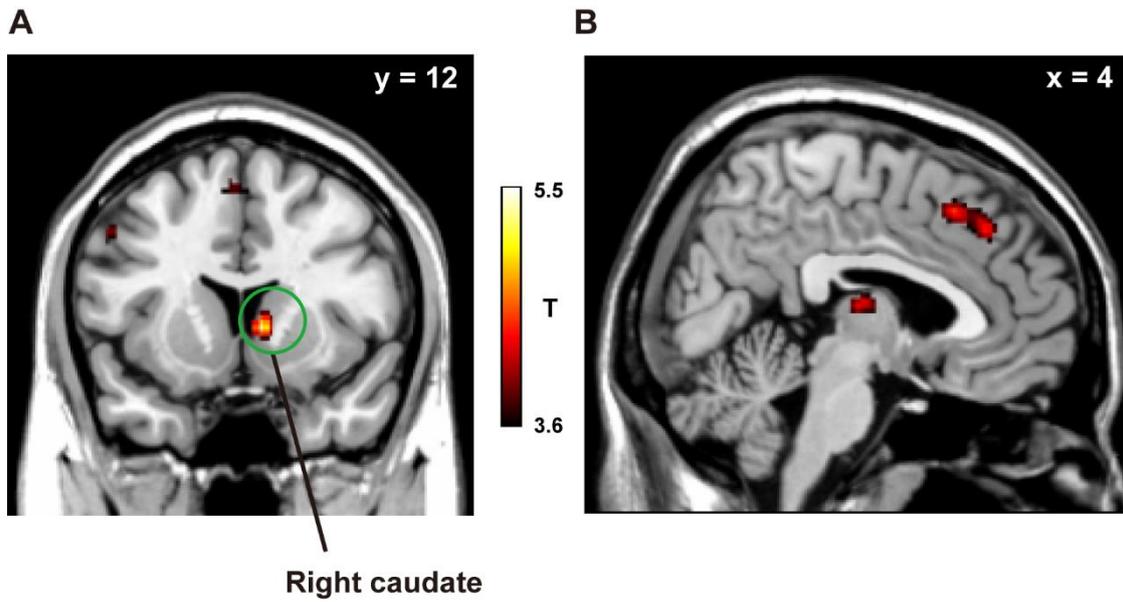
Figure 17. Activation modulated by SA only for the correct trials. Neural activations parametrically modulated by SA, only for the correct trials. The cortical activation map was projected onto the coronal (A) and sagittal (B) plains (uncorrected $p < 0.001$). See Table 4 for the stereotactic coordinates.

Brain region	BA	Side	x	y	z	Z-Value	Voxels
dmPFC	32	M	6	34	40	3.8	181
	8	M	4	22	48	3.6	*
vmPFC	11	R	26	38	-14	3.6	14
dlPFC	46	R	40	54	2	3.9	82
vlPFC	47	R	28	32	-4	3.9	71
LPMC	6	R	50	4	50	3.5	71
			38	2	56	3.3	*
SMA	6	M	-4	10	56	3.4	41
IFG	44	L	-52	12	40	3.3	22
	44	R	52	10	46	3.1	1
	45	R	44	32	34	3.6	121
			46	32	22	3.3	*
Insula	13	L	-24	28	4	4.1	15
			-34	16	8	3.8	99
Insula		R	42	18	2	3.5	35
Caudate		R	8	12	2	4.2	77
Thalamus		M	4	-16	10	3.5	53
		M	-8	-4	-2	3.2	5
PG	30	L	-20	-30	-10	3.3	3
ACC	24	M	12	36	22	3.4	10
MCC	23	M	0	-12	30	3.2	1
Precuneus	7	L	-14	-68	34	3.7	22
IPL	40	R	34	-38	36	3.3	1
LOG	19	R	24	-58	34	3.1	1
			28	-80	4	3.5	67
MOG	18	R	18	-90	0	3.5	*
		M	-8	-72	-24	3.3	6
OP	17	R	28	-96	10	3.4	41
Cerebellum		L	-36	-38	-36	3.7	27

Table 4. Regions where activations were modulated by SA for the correct trials.
Stereotactic coordinates (*x, y, z*) in the MNI space (mm) are shown for each activation peak of Z-values. MCC, middle cingulate cortex; SMA, supplementary motor area; dmPFC, dorsomedial prefrontal cortex; PG, parahippocampal gyrus.
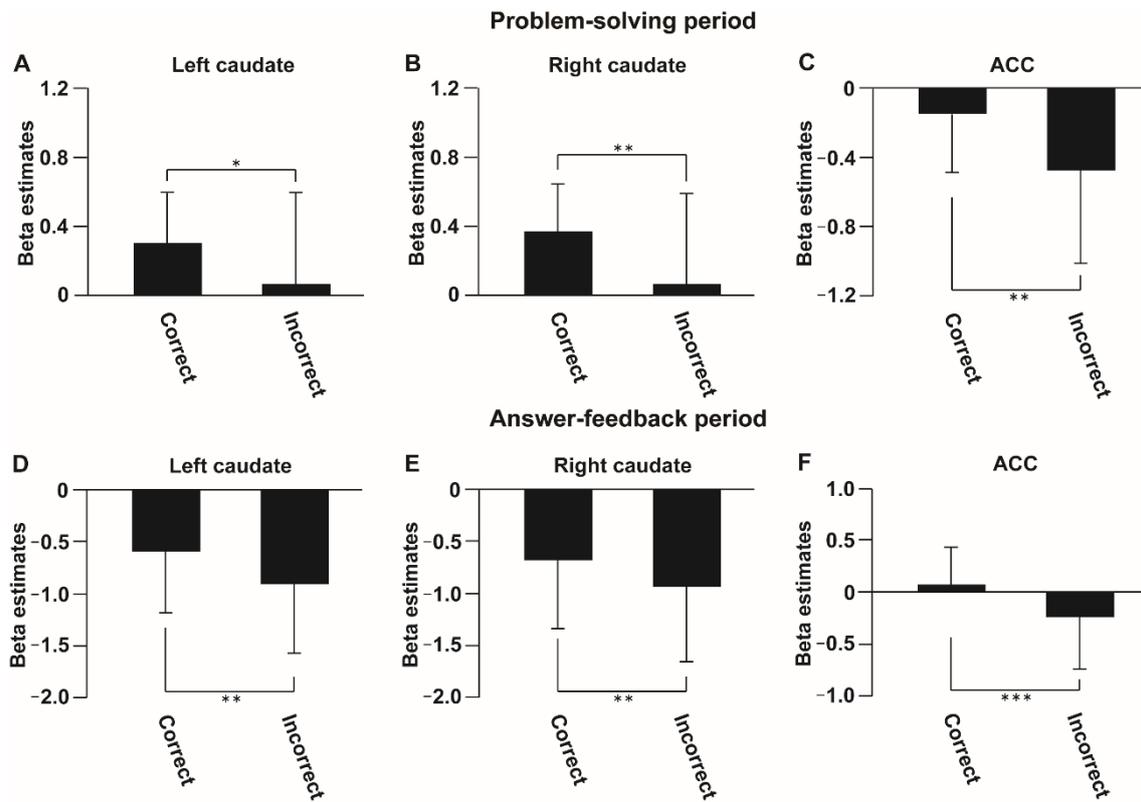
Figure 18. Activations in the correct and incorrect trials. Beta estimates of correct and incorrect trials were extracted from the anatomically defined ROIs of the left caudate (A, D), right caudate (B, E), and ACC (C, F) for both problem-solving period and answer-feedback period. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$. Error bars, SD.

4.4. Discussion

In the present experiment, I prepared three levels of instruction and a basic control for the same insight-problem task with number and letter notations, and I obtained following results. First, I found that the Indirect hint induced the largest SA among the three instruction levels. I also confirmed that the Indirect hint corresponded well to the intermediate level of difficulty for each participant. Secondly, using fMRI I found activations in the left IFG and left LPMC during problem-solving period, which reflect the processing of both number and letter sequences. Thirdly, I observed that activations in the bilateral caudate, together with the ACC, PCC, vmPFC, and vlPFC, showed significant modulation with SA during the problem-solving period. Fourthly, in the bilateral caudate the intermediate level difficulty trials induced the largest activations. Lastly, during answer-feedback period, the activation in the ACC was modulated by SA, while the ACC activation seemed to reflect differences between the correct and incorrect trials. These results indicate that the Indirect instruction can most effectively contribute to the learners' SA, and that such feeling is principally processing by the bilateral caudate in the reward system of the brain.

Compared to the task design of Experiment 3, the task design in Experiment 4 does not explicitly require calculation or processing of sentences. However, cumulative addition is necessary to report the fourth term of a sequence "3, 5, 7, …". Indeed, I revealed that the calculation of linear sequences or quadratic sequences induce activations in the left IFG (Nakai & Sakai, 2014). In this study I analyzed the calculation processes behind linear/quadratic sequences, and showed tree structure descriptions. Therefore, the activation in the left IFG in Experiment 4 probably reflect syntactic processing related to the calculation of sequences. As for the language task, I used Hiragana which is a symbol sequence system. Although Hiragana sequence does not require calculation, the computation process can be described by the hierarchical tree structures exactly in the same manner as the number sequence. In the study of Nakai and Sakai (2014), authors showed that activations in the left IFG is independent of the number of operations (e.g., addition and subtraction), by comparing two task with the same number of operations, but with different tree structures. A transition from certain Hiragana "KA" to the next one "KI" (or after the second Hiragana "KU") is comparable to the addition +1 (or +2). In Experiment 4, I did not directly analyze functional connectivity between reward system

and left IFG. The involvement of the caudate nucleus is suggested in artificial grammar learning (Chan et al., 2013; Forkstam et al., 2006; Lieberman et al., 2004). The next step would be to reveal a relationship between the left IFG and caudate reward system during learning period of language/mathematics.

Many fMRI studies have observed an increase in activations, specifically in the frontal and parietal regions, in tasks with high cognitive demands compared to those with low demands, across various cognitive domains (Fedorenko, Duncan, & Kanwisher, 2013; Newman, Carpenter, Varma, & Just, 2003). Indeed, I found activation in the left IFG, left LPMC, and bilateral parietal cortex when the participants solved problems compared to the Control condition. However, I also showed that higher task difficulty does not always induce larger activations in reward systems. This view is supported by a recent neuroimaging study, which revealed that both cognitive effort and motor effort modulated activations in the bilateral striatum, independent of task difficulty (Schmidt, Lebreton, Cléry-Melin, Daunizeau, & Pessiglione, 2012). Although two notations (Number and Letter) had different intrinsic problem difficulties, the effect of the Indirect condition was robust. The current results show that the reward and emotion-related regions contribute to SA in a domain-general manner, most likely playing a complementary role with front-parietal networks during problem solving.

Research shows that tasks with intermediate difficulty induce the greatest amount of pleasure after solving the task (Harter, 1978a, 1978b). The current results are consistent with this view, since the Indirect hint successfully imposed an intermediate difficulty level, and induced the largest SA among the three instruction levels. Furthermore, another previous study also indicated that the strength of motivation to perform a task was maximized with intermediate difficulty and with the largest uncertainty of success (Atkinson, 1957). Even though in the correct trials Indirect condition does not provide larger SA under the Indirect condition than under the No hint condition, the percentage of correct trials (i.e., accuracy) under the No hint trials is smaller than that under the Indirect condition, which cancels out the superiority of No hint condition for the overall SA. The current results provide new evidence to bridge the gap between task difficulty, SA, and motivation for learning. I should note, however, the limitation of the above interpretation of the link between SA and motivation. I assumed that SA value in the correct and incorrect trials had an equal weight, i.e., SA value "3" in the No hint condition

provides the same amount of motivation as SA value "3" in the Direct condition. I averaged SA in the correct and incorrect trials without considering such possibility. Further is necessary to confirm the link between SA and subsequent behavioral changes.

The bilateral caudate has been reportedly involved in reward-based learning (Haruno et al., 2004), as well as in motivation studies (Delgado et al., 2004; Murayama et al., 2010). Other imaging study also suggests involvement of caudate in the artificial grammar learning of letter sequences (Lieberman et al., 2004). Although the authors did not discuss the involvement of caudate in the reward-processing, caudate activation in the contrast of grammatical vs. nongrammatical sequences may reflect the SA with successful grammatical judgments. In the current experiment, I found that the bilateral caudate principally reflected SA during problem-solving period, which was mirroring behavioural results of the superiority of the Indirect hint, supporting the assumption that the caudate is involved in motivation-enhancing aspect in learning. It is reported that dorsal striatum (including caudate) is involved in the maintenance of information about the rewarding outcomes necessary for learning, while ventral striatum is simply involved in the anticipation of future reward (O'Doherty et al., 2004). The current results suggest that the sense of accomplishment is processed rather in the caudate which has particular importance in motivation for learning.

Studies have shown the activation of the ACC is related to processing emotional saliency (Goldstein et al., 2009; Phan et al., 2004). The current experimental results indicate that the reward system interacts with this emotional processing system. Although activation cluster in the ACC seems to have an overlap with the corpus callosum, I confirmed that the peak voxel in the cluster is included in the ACC. It is likely that the smoothing method adopted in the analysis procedure may have extended the activation cluster. Furthermore, activation of the PCC was observed during processing emotional words (Maddock, Garrett, & Buonocore, 2003), and emotion-based moral judgments (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). I found both positive and negative beta estimates for the signal changes data. The reference point of zero in those data does not reflect the activation in resting states, but is an average signal of the target volume from all time points included in the analysis, provided that session effects are regressed out. One possible reason of the negative beta estimates in the current experiment is that I did not include resting condition in the experimental design. In such

case about a half of time points would show a negative value. The crucial question is not whether the signals are positive or negative, but whether the signals increased (activation) or decreased (deactivation) compared to the reference point. As for the negative activations in the ACC during problem-solving period, it is possible that this negative BOLD is due to the task-specific deactivation of the default mode network (Raichle, 2013). Task-induced deactivation in the ACC was also reported in the domain of motor response (Kudo et al., 2004). Considering the modulation by SA, such baseline deactivation may be further affected by emotional saliency.

During the answer-feedback period, I found modulation with SA in the ACC/vmPFC, but not in the bilateral caudate, which is consistent with a previous neuroimaging study which reported activation in the ACC/vmPFC for reward outcomes (Knutson, Fong, Adams, Varner, & Hommer, 2001). I did not find modulation of activations in the ACC when I analysed only with the correct trials. It is conceivable that the ACC activation may be induced by conflict resolution (i.e., knowing the correct answer to the given problems), in such case the incorrect trials are with "higher conflict state" and may induce larger activation. However, activations in the ACC were larger in correct trials than in incorrect trials, even during answer-feedback period. This result cannot be explained by the conflict-resolution interpretation, suggesting that the ACC likely reflects the differential emotional saliency between correct and incorrect trials rather than SA itself.

By dividing all trials into groups according to the length of the RT, I considered individual variability in their skills. I found that trials with intermediate RTs induced the largest SA, and the largest activations in the bilateral caudate. Since indirect instructions correspond well to the trials with intermediate RTs, indirect instructions may be effective regardless of individual skill. It is possible that some individuals prefer to solve problems by themselves, whereas others prefer to be instructed step by step. Therefore, it remains important to adjust problem difficulty and instruction type to each individual to obtain an optimal efficiency for problem solving. Moreover, based on the expectancy-value theory, the effect of problem difficulty on the attitude toward success may depend on the personality trait of motivation to achieve success and avoidance of failure (Atkinson, 1964). It is possible that some individuals prefer to solve problems by themselves, whereas others prefer to be instructed step by step. It is likely that for individuals who prefer instructions, giving instructions itself can be viewed as a reward. Therefore, it

remains important to adjust problem difficulty and instruction type to each individual to obtain an optimal efficiency for problem solving. Although it may be problematic to identify SA and the attitude toward success, a future study may clarify the relationship between individual variability of motivation toward learning and SA/activation in the reward system.

It would be important to discuss some technical limitations in the current experimental setting. First, although in the problem-solving period the trial onsets were modeled based on the participants' response timing and were automatically jittered in each trial (with a low correlation coefficient value of regressors shown in Supplement texts), in the answer-feedback period the trial onsets were fixed. This experimental setting may lead to the underestimation of the effects in the answer-feedback period. Secondly, for the sake of fMRI analysis based on the participants' button presses, I instructed participants to press any button when they couldn't find an answer. Under this experimental design I could not distinguish between participants' random guess and the true insight. However, those forced responses would result in slow RT and thus included in Hard trials, and in such unsuccessful trials SA would be small. In the current paradigm the Medium trials are the most important component for considering the effect of SA on caudate. Therefore the negative effect of this instruction was limited.

The application of neuroscience to education, called "Neuroeducation", has recently attracted a wide variety of attention (Ansari, De Smedt, & Grabner, 2012; Carew & Magsamen, 2010). There has been many discussions regarding the level of instruction that should be provided for students (Kirschner, Sweller, & Clark, 2006; Lee & Anderson, 2013; Reiser, Copen, Ranney, Hamid, & Kimberg, 1998). Proper instructions may change the subjective difficulty, even within the same problem, thereby encouraging students to be successful in problem solving. One such putative instruction strategy is the discovery learning strategy (Bruner, 1961; Reiser et al., 1998), in which minimal instruction may encourage students to discover knowledge by themselves, rendering them intrinsically motivated to learn. Conversely, several studies have claimed that direct instruction, which provides sufficient information regarding concepts and procedures necessary for problem solving, is more efficient for learning (Craig, 1956; Kirschner, Sweller, & Clark, 2006). The current results indicate that both extremes are ineffective, and that the Indirect hint has a positive influence on the participants' SA. Although the participant obtained highly

accurate performance with direct instructions, it seemed that they did not report satisfaction because the tasks had become too easy. Similarly, the No hint condition also induced a reduced SA in comparison to the Indirect condition, mirroring the lower success rate under this condition. Under the Indirect condition, the participants were given partial information regarding problem solving. Such instruction allows the participant to discover information by himself or herself, considered an example of "guided discovery" (Lee & Anderson, 2013; Mayer, 2004). Our current results are suggestive for the field of Neuroeducation, by giving a neuroscientific basis on instructional methods for students. I conclude that reported sense of accomplishment is effectively induced by the Indirect instruction method, and represented in the brain's reward system. The present results will provide new insights on the sense of accomplishment and brain reward system, which may support the problem-solving of linguistic and numerical tasks.

Chapter 6. General Discussion

6.1. Summary of the four experimental results

I performed series of four experiment to reveal shared neural basis for linguistic and numerical tasks, and found the following novel evidence. In the Experiment 1 (Chapter 2), I measured GABA concentration in the bilateral IFG using H^1-MRS, and found the correlation of GABA concentration and language performance (categorical fluency scores) in the human brain for the first time. In the Experiment 2 (Chapter 3), I newly created a cross-domain structural priming paradigm between linguistic and numerical tasks, and found priming effect of error rates only with participants in the natural science departments. In the Experiment 4 (Chapter 4), I showed the cross-domain structural interaction (RS effect) between linguistic and numerical tasks in the left IFG for the first time. In the Experiment 4 (Chapter 5), I again found activation in the left IFG for the linguistic and numerical tasks, and also found that activations in the bilateral caudate nucleus were modulated by the sense of accomplishment values related to the linguistic and numerical tasks.

6.2. Specific features of the left IFG

The left IFG, traditionally called Broca's area, is considered as a critical region for language processing. In the series of experiments I sought to reveal the functional specificity and neurotransmitter basis of the IFG using linguistic and numerical tasks. Studies showed that anatomical basis of Broca's area (Amunts et al., 1999; Amunts et al., 2003; Foundas et al., 1996). The first experiment aimed at providing evidence of neurotransmitter basis of this region. No significant difference of GABA concentration was found between the left IFG and right IFG, while I found a correlation between verbal fluency scores and GABA+/Cr ratio in the left IFG. Studies showed negative correlation of GABA concentration and BOLD signal during resting state in the visual area (Muthukumaraswamy et al., 2009), as well as in the ACC (Northoff et al., 2007). The results of the experiment 1 indicate a link between the neurotransmitter concentration and language function in the left IFG for the first time.

It is recently reported that the left IFG is also important for numerical cognition (Baldo & Dronkers, 2007; Makuuchi et al., 2012; Nakai & Sakai, 2014), but there are

also opposite evidences (Amalric & Dehaene, 2016; Ansari, 2016; Fedorenko et al., 2011; Maruyama et al., 2012; Monti et al., 2012). To clearly assess the functional specificity of the left IFG, it is necessary to examine relationship between different cognitive domains associated with this region. The experiment 2, 3, and 4 were aimed at providing evidence of functional commonality of the left IFG for the linguistic and numerical tasks.

The major concern against the idea of shared syntactic processing for the linguistic and numerical tasks in the left IFG is that (1) working memory is also a candidate for explain the role of IFG, (2) the activation overlaps in the left IFG does not ensure the same neural circuits in the same region, and (3) syntactic structures in m-expressions were confounded with the effect of operators. To answer these issue, I prepared cross-domain structural priming paradigm in experiment 2, and confirmed the priming effect for students in natural science department. In experiment 3, I used this design with the fMRI measurement. The RS effect in the left IFG indicates that activation changes in the left IFG is independent of working memory, because both congruent and incongruent pairs had the same working memory load. The experiment 3 further showed indicates that activations in the left IFG for the linguistic and numerical tasks are based on the same neural circuit, because the RS effect between two stimuli may indicate changes of activation patterns in a certain neural circuit (Grill-Spector, Henson, & Martin, 2006).

In the Experiment 4 I confirmed activation in the left IFG for a numerical task and corresponding linguistic task. The previous study of Nakai and Sakai (2014) has reported activation in the left IFG for the numerical task with recursive computation using linear sequence such as "1, 3, 5, …", but was criticized that they failed to control the number of operations (addition and subtraction) (Hung et al., 2015). The result of the Experiment 4 exclude this possibility, because I did not use any operations with the letter notation. The results of experiment 3 also suggest that activation changes in the left IFG is independent of operators, because m-expressions in the congruent and incongruent conditions had the same number of operators. The results of the experiment 2 to 4 repeatedly indicate the commonality of the left IFG for linguistic and numerical tasks,

which can be regarded as useful information for preparing future research on the neurotransmitter basis of mathematics.

6.3. Multi-domain functions of Broca's area

Previous studies suggest that syntactic structures are found not only in m-expressions, but also in music (Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Maess, Koelsch, Gunter, & Friederici, 2001; Rohrmeier, 2011), and action (Fazio et al., 2009). Interestingly, some researchers even try to explain the theory of mind in terms of recursive computation (de Villiers, 2007; Jenkins & Astington, 1996). According to Rohrmeier (2011), dominant-tonic motion such as "G7" to "Am" is a central structure of the context-free grammar description of the western tonal harmony. Neuroimaging study showed activation in the right IFG reflect the detection of syntactic violation of tonal harmony (Maess et al., 2001). It has been proposed that the recursive computation is the only basis of human-unique component of language faculty (Hauser et al., 2002). This very strong hypothesis has been criticized by several researchers (Pinker & Jackendoff, 2005). It seems that the recursion-only hypothesis is theoretically implausible at least concerning numerical representation, where the abstract symbol representation may not need syntactic processing, but it is clearly human-unique. Interactions between recursion and other cognitive factors would be important for producing human-unique capacities.

6.4. Limitations of current experiments

It is worth noting the limitation of current experiments. First, I extracted only small components of language and mathematics. At least for mathematics there are many components which should be taken in account, such as abstract algebra, geometry, and logical inference. For example, I ignored semantics in m-expressions. Here I tentatively assume the semantics of mathematics as external references of mathematical symbols (note that this is an extensional definition of semantics). Compared to the m-expression such as "3+5×2", an abstract m-expression "x+y×z" may lack the semantic component. Operators such as + and × can be considered as functional words according to a previous study (Makuuchi et al., 2012). Neuroimaging studies with jabberwocky sentences (sentences with pseudo words) also showed activations in the left IFG (Ohta, Fukui, & Sakai, 2013; Pallier et al., 2011). Therefore, it is naturally predicted that activations in the

left IFG is also induced with abstract algebra without number symbols. In contrast, numerical processing in the bilateral IPL may be considered as semantic component. Indeed, the contrast of congruent and incongruent pairs in Experiment 3 excluded effect of semantic by comparing stimuli pairs with exactly same semantic contents. The left IFG activations in experiment 4 reflected only syntactic component. Further studies, however, are necessary to segregate semantic processing from syntactic processing in arithmetic.

Secondly, in the experiment 2 I found significant cross-domain priming effect in the error rates data, for both Lang to Num and Num to Lang directions. Interestingly, the priming effect was found for the students in natural science department. This result may suggest that structural sensitivity of m-expressions depend on the experience or knowledge of individuals. It might be argued that such feature is inconsistent with the idea of the shared neural basis of linguistic and numerical tasks, because the commonality of the two cognitive functions is not open for everyone. No kids can solve m-expressions without education. It is likely that structural sensitivity to m-expressions is elaborated by education or experience in mathematics.

Thirdly, although I recruited both students in both natural science and humanity/social science departments in Experiment 2 (behavioral experiment), I recruited only students in natural science department in Experiment 3 (fMRI experiment). I did not even investigate the department of participants in Experiment 1 and 4. The natural science experience of participants may not be relevant to the verbal fluency scores in Experiment 1. The scientific experience may affect the task difficulty in Experiment 4, but I also revealed that the participants felt the largest sense of accomplishment for the problems with intermediate difficulty for themselves. Thus, the result of Experiment 4 is applicable even for students in natural science departments. Research showed different functional/anatomical property in the brain for expert mathematicians (Amalric & Dehaene, 2016; Aydin et al., 2007). Similarly to the priming effect in Experiment 2, the repetition suppression effect in the left IFG in Experiment 3 may also differ depending on the expertise in mathematics. In the current experiment, I confirmed that repetition suppression effect is found in the left IFG. The next step would be to investigate individual variability of such effect.

Lastly, in the Experiment 4, I did not find a significant difference in the fMRI data in a certain contrast, while I found a significant difference in behavioral data. For example,

the contrast of Direct vs. Control did not show significant activations, which seems contradictory to the statement that fMRI measurement is more sensitive than behavioral measurement. However, I confirmed a significant difference between Direct and Control conditions in the signal changes extracted in the left IFG and LPMC. The whole-brain analysis is sometimes too conservative to detect a meaningful differences. In such case extracted signals in anatomical masks can be used for the detailed comparisons. Furthermore, in the experiment 3 I used the multi-band EPI sequence, which has better spatial and temporal resolution than normal EPI used in experiment 4. This technical advantage might provide more sensitive results.

6.5. Conclusions

In the series of experiments, I aimed at examining the functional specificity and neurotransmitter basis of the IFG using linguistic and numerical tasks. I concluded that syntactic structures of linguistic and numerical tasks are processed in the opercular part of the left IFG (the posterior portion of Broca's area), which has a specific feature of neurotransmitter reflecting individual variability of language performance, while sense of accomplishment related to the linguistic and numerical tasks is processed in the bilateral caudate nucleus.

References

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences*, *113*(18), 4909–4917.

Amunts, K., Schleicher, A., Ditterich, A., & Zilles, K. (2003). Broca's region: cytoarchitectonic asymmetry and developmental changes. *The Journal of Comparative Neurology*, *465*(1), 72–89.

Amunts, K., Schleicher, a, Bürgel, U., Mohlberg, H., Uylings, H. B., & Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *The Journal of Comparative Neurology*, *412*(2), 319–41.

Amunts, K., Weiss, P. H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J. M., … Zilles, K. (2004). Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space - The roles of Brodmann areas 44 and 45. *NeuroImage*, *22*(1), 42–56.

Ansari, D. (2016). The neural roots of mathematical expertise. *Proceedings of the National Academy of Sciences*, *113*(18), 4887–4889.

Ansari, D., De Smedt, B., & Grabner, R. H. (2012). Neuroeducation - A critical overview of an emerging field. *Neuroethics*, *5*(2), 105–117.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*(6), 359–372.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Aydin, K., Ucar, A., Oguz, K. K., Okur, O. O., Agayev, A., Unal, Z., … Ozturk, C. (2007). Increased gray matter density in the parietal cortex of mathematicians: A voxel-based morphometry study. *American Journal of Neuroradiology*, *28*(10), 1859–1864.

Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*, *42*(2), 525–534.

Baldo, J. V, & Dronkers, N. F. (2007). Neural correlates of arithmetic and language comprehension: a common substrate? *Neuropsychologia*, *45*(2), 229–35.

Basso, A., Burgio, F., & Caporali, A. (2000). Acalculia, Aphasia and Spatial Disorders

in Left and Right Brain-Damaged Patients. *Cortex*, *36*(2), 265–280.

Basso, A., Caporali, A., & Faglioni, P. (2005). Spontaneous recovery from acalculia. *Journal of the International Neuropsychological Society : JINS*, *11*(1), 99–107.

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia*, *6*(1), 53–60.

Boy, F., Evans, C. J., Edden, R. a E., Singh, K. D., Husain, M., & Sumner, P. (2010). Individual differences in subconscious motor control predicted by GABA concentration in SMA. *Current Biology : CB*, *20*(19), 1779–85.

Broca, P. (1861a). Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bulletin de La Société Anthropologique*, *2*, 235–238.

Broca, P. (1861b). Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin de La Société Anatomique*, *6*, 330–357.

Bruner, J. (1961). The Act of Discovery. *Harvard Educational Review*. Harvard educational review.

Cantalupo, C., & Hopkins, W. D. (2001). Asymmetric Broca's area in great apes. *Nature*, 414(6863), 505–505.

Carew, T. J., & Magsamen, S. H. (2010). Neuroscience and Education: An Ideal Partnership for Producing Evidence-Based Solutions to Guide 21st Century Learning. *Neuron*, *67*(5), 685–688.

Chan, S. H., Ryan, L., & Bever, T. G. (2013). Role of the striatum in language: Syntactic and conceptual sequencing. *Brain and Language*, *125*(3), 283–294.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (2006). *Language and mind*. Cambridge University Press.

Chumbley, J. R., & Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage*, *44*(1), 62–70.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum;

Costafreda, S. G., Fu, C. H. Y., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. *Human Brain Mapping*, *27*(10),

799–810.

Craig, R. C. (1956). Directed versus independent discovery of established relations. *Journal of Educational Psychology*, *47*(4), 223–234.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron*, *24*(2), 427–32.

de Villiers, J. (2007). The Interface of Language and Theory of Mind. *Lingua. International Review of General Linguistics. Revue Internationale de Linguistique Generale*, *117*(11), 1858–1878.

Dehaene, S., Spelke, E. S., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain Imaging Evidence. *Science*, *284*(5416), 970–974.

Dejerine, J. (1891). Sur un cas de cécité verbale avec agraphie suivi d'autopsie. *Mémoires de La Société de Biologie*, *3*, 197–201.

Delazer, M., Girelli, L., Semenza, C., & Denes, G. (1999). Numerical skills and aphasia. *Journal of the International Neuropsychological Society*, *5*(3), 213–21.

Delgado, M. R., Stenger, V. a., & Fiez, J. a. (2004). Motivation-dependent responses in the human caudate nucleus. *Cerebral Cortex*, *14*(9), 1022–1030.

Devauchelle, A.-D., Oppenheim, C., Rizzi, L., Dehaene, S., & Pallier, C. (2009). Sentence syntax and content in the human temporal lobe: an fMRI adaptation study in auditory and visual modalities. *Journal of Cognitive Neuroscience*, *21*(5), 1000–1012.

Devlin, K. J. (2000). *The math gene: How mathematical thinking evolved and why numbers are like gossip.* New York: Basic Books.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, *130*(5), 1432–1441.

Dunbar, R. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*, 681–694.

Edden, R. A. E., Muthukumaraswamy, S. D., Freeman, T. C. A., & Singh, K. D. (2009). Orientation discrimination performance is predicted by GABA concentration and gamma oscillation frequency in human primary visual cortex. *Journal of Neuroscience*, *29*(50), 15721–15726.

Edden, R. A. E., Puts, N. A. J., Harris, A. D., Barker, P. B., & Evans, C. J. (2014). Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *Journal of Magnetic Resonance Imaging*, *40*(6), 1445–1452.

Ernest, P. (1987). a Model of the Cognitive Meaning of Mathematical Expressions. *British Journal of Educational Psychology*, *57*(3), 343–370.

Exner, S. (1881). *Untersuchungen über die Localisation der Functionen in der Grosshirnrinde des Menschen*. Wien: W Braumuller.

Fazio, P., Cantagallo, A., Craighero, L., D'ausilio, A., Roy, A. C., Pozzo, T., … Fadiga, L. (2009). Encoding of human action in Broca's area. *Brain*, *132*(7), 1980–1988.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428–16433.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, 22(21), 2059-2062.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16616–16621.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177–94.

Ford, A. a, Triplett, W., Sudhyadhom, A., Gullett, J., McGregor, K., Fitzgerald, D. B., … Crosson, B. (2013). Broca's area and its striatal and thalamic connections: a diffusion-MRI tractography study. *Frontiers in Neuroanatomy*, *7*, 8.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage*, *32*(2), 956–967.

Foundas, A. L., Eure, K. F., Luevano, L. F., & Weinberger, D. R. (1998). MRI asymmetries of Broca's area: the pars triangularis and pars opercularis. *Brain and Language*, *64*(3), 282–296.

Foundas, A. L., Leonard, C. M., Gilmore, R. L., Fennell, E. B., & Heilman, K. M.

(1996). Pars triangularis asymmetry and language dominance. *Proceedings of the National Academy of Sciences*, *93*(2), 719–722.

Friederici, A. D., Bahlmann, J., Friedrich, R., & Makuuchi, M. (2011). The Neural Basis of Recursion and Complex Syntactic Hierarchy. *Biolinguistics*, *5*(1), 87–104.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2458–63.

Geschwind, N. (1965). Disconnexion syndromes in animal and man. *Brain*, *88*, 237–294.

Geschwind, N. (1970). The Organization of Language and the Brain: Language disorders after brain damage help in elucidating the neural basis of verbal behavior. *Science*, *170*(3961), 940–944.

Glick, S. D., Ross, D. A., & Hough, L. B. (1982). Lateral asymmetry of neurotransmitters in human brain. *Brain Research*, *234*(1), 53–63.

Goldstein, R. Z., Alia-Klein, N., Tomasi, D., Carrillo, J. H., Maloney, T., Woicik, P. A., … Volkow, N. D. (2009). Anterior cingulate cortex hypoactivations to an emotionally salient task in cocaine addiction. *Proc Natl Acad Sci U S A*, *106*(23), 9453–9458.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, *12*(12), 474–480.

Harter, S. (1978a). Effectance Motivation Reconsidered Toward a Developmental Model. *Human Development*, *21*(1), 34–64.

Harter, S. (1978b). Pleasure Derived from Challenge and the Effects of Receiving Grades on Children's Difficulty Level Choices. *Child Development*, *49*(3), 788–799.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., … Kawato,

M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *The Journal of Neuroscience* , *24*(7), 1660–1665.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–79.

Heim, S., Opitz, B., Müller, K., & Friederici, A. D. (2003). Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Cognitive Brain Research*, *16*(2), 285–296.

Homan, P., Vermathen, P., Van Swam, C., Federspiel, A., Boesch, C., Strik, W., … Kreis, R. (2014). Magnetic resonance spectroscopy investigations of functionally defined language areas in schizophrenia patients with and without auditory hallucinations. *NeuroImage*, *94*, 23–32.

Hopcroft, J., Motwani, R., & Ullman, J. (2006). *Automata theory, languages, and computation, 3rd Edition,* Addison-Wesley.

Hopkins, W. D., Taglialatela, J. P., Meguerditchian, A., Nir, T., Natalie, M., Sherwood, C. C., & College, A. S. (2009). Gray Matter Asymmetries in Chimpanzees as Revealed by Voxel- Based Morphometry. *NeuroImage*, *42*(2), 491–497.

Hung, Y. H., Pallier, C., Dehaene, S., Lin, Y. C., Chang, A., Tzeng, O. J. L., & Wu, D. H. (2015). Neural correlates of merging number words. *NeuroImage*, *122*, 33–43.

Hurford, J. R. (1987). *Language and number: The emergence of a cognitive system.* Basil Blackwell, Oxford.

Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., & Delazer, M. (2007). Imaging early practice effects in arithmetic. *NeuroImage*, *36*(3), 993–1003.

Izard, V., Pica, P., Spelke, E., & Dehaene, S. (2008). Exact Equality and Successor Function: Two Key Concepts on the Path towards understanding Exact Numbers. *Philosophical Psychology*, *21*(4), 491–505.

Jansen, A. R., Marriott, K., & Yelland, G. W. (2003). Comprehension of algebraic expressions by experienced users of mathematics. *The Quarterly Journal of Experimental Psychology*, *56A*(1), 3–30.

Jansen, A. R., Yelland, G. W., & Marriott, K. (2007). Parsing of Algebraic Expressions by Experienced Users of Mathematics Corresponding Author : *European Journal of Cognitive Psychology*, *19*(2), 286–320.

Jenkins, J., & Astington, J. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, (1), 70–78.

Joshi, A. K. (1985). Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In D. Dowty, L. Kartunnen, A. Zwickey (Eds), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*. Cambridge Unversity Press.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., … Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLoS Biology*, *2*(4), E97.

Kansaku, K., Johnson, A., Grillon, M.-L., Garraux, G., Sadato, N., & Hallett, M. (2006). Neural correlates of counting of sequential sensory and motor events in the human brain. *NeuroImage*, *31*(2), 649–60.

Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience*, *1*(5), 411–416.

Keller, J. M. (1987). Strategies for stimulating the motivation to learn. *Performance + Instruction*, *26*(8), 1–7.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During *Educational Psychologist*, *41*(2), 75–86.

Kirshner, D. (1989). The Visual Syntax of Algebra. *Journal for Research in Mathematics Education*, *20*(3), 274–287.

Klessinger, N., Szczerbinski, M., & Varley, R. (2007). Algebra in a man with severe aphasia. *Neuropsychologia*, *45*(8), 1642–8.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*(17), 3683–3687.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), 15443–8.

Krueger, F., Spampinato, M. V, Pardini, M., Pajevic, S., Wood, J. N., Weiss, G. H., … Grafman, J. (2008). Integral calculus problem solving: an fMRI investigation. *Neuroreport*, *19*(11), 1095–1099.

Kudo, K., Miyazaki, M., Kimura, T., Yamanaka, K., Kadota, H., Hirashima, M., ... & Ohtsuki, T. (2004). Selective activation and deactivation of the human brain structures between speeded and precisely timed tapping responses to identical visual stimulus: an fMRI study. *NeuroImage*, 22(3), 1291-1301.

Lee, H. S., & Anderson, J. R. (2013). Student Learning: What Has Instruction Got to Do With It? *Annual Review of Psychology*, *64*(1), 445–469.

Lepper, M. R., & Hodell., M. (1989). Intrinsic motivation in the classroom. *Research on Motivation in Education*, *3*, 73–105.

Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., & Knowlton, B. J. (2004). An Event-related fMRI Study of Artificial Grammar Learning in a Balanced Chunk Strength Design. *J Cogn Neurosci*, *16*(3), 427–438.

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*, *4*(4), 423–428.

Linchevski, L., & Livneh, D. (1999). Structure sense: The relationship between algebraic and numerical contexts. *Educational Studies in Mathematics*, *40*(2), 173–196.

Maddock, R. J., Garrett, A. S., & Buonocore, M. H. (2003). Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Human Brain Mapping*, *18*(1), 30–41.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, *4*(5), 540–545.

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(20), 8362–7.

Makuuchi, M., Bahlmann, J., & Friederici, A. D. (2012). An approach to separating the levels of hierarchical structure building in language and mathematics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 2033–2045.

Weiss, M. R. (2000). Motivating kids in physical activity. *President's Council on Physical Fitness and Sports Research Digest, 3*(11), 1–8.

Martín-Loeches, M., Casado, P., Gonzalo, R., De Heras, L., & Fernández-Frías, C. (2006). Brain potentials to mathematical syntax problems. *Psychophysiology*, *43*(6), 579–91.

Maruyama, M., Pallier, C., Jobert, A., Sigman, M., & Dehaene, S. (2012). The cortical representation of simple mathematical expressions. *NeuroImage*, *61*(4), 1444–60.

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? The case for guided methods of instruction. *American Psychologist*, *59*(1), 14–19.

Menon, V., Mackenzie, K., Rivera, S. M., & Reiss, A. L. (2002). Prefrontal Cortex Involvement in Processing Incorrect Arithmetic Equations : Evidence From Event-Related fMR. *Human brain mapping, 16(2), 119-130.*

Mescher, M., Merkle, H., Kirsch, J., Garwood, M., & Gruetter, R. (1998). Simultaneous in vivo spectral editing and water suppression. *NMR in Biomedicine*, *11*(6), 266–272.

Michels, L., Martin, E., Klaver, P., Edden, R., Zelaya, F., Lythgoe, D., … O'Gorman, R. (2012). Frontal gaba levels change during working memory. *PLOS ONE*, *7*(4), e31933.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.), *The Frontal Granular Cortex and Behavior* (pp. 313–334). New York: McGraw Hill.

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain FMRI. *Magnetic Resonance in Medicine*, *63*(5), 1144–1153.

Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought beyond language: neural dissociation of algebra and natural language. *Psychological Science*, *23*(8), 914–22.

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(49), 20911–6.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., & Weiller,

C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, *6*(7), 774–81.

Muthukumaraswamy, S. D., Edden, R. a E., Jones, D. K., Swettenham, J. B., & Singh, K. D. (2009). Resting GABA concentration predicts peak gamma frequency and fMRI amplitude in response to visual stimulation in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(20), 8356–61.

Nakai, T., & Sakai, K. L. (2014). Neural Mechanisms Underlying the Computation of Hierarchical Tree Structures in Mathematics. *PLOS One*, *9*(11), e111439.

Newman, S. D., Carpenter, P. A., Varma, S., & Just, M. A. (2003). Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia*, *41*(12), 1668–1682.

Newman, S. D., & Twieg, D. (2001). Differences in auditory processing of words and pseudowords: An fMRI study. *Human Brain Mapping*, *14*(1), 39–47.

Noppeney, U., & Price, C. J. (2004). An FMRI study of syntactic adaptation. *Journal of Cognitive Neuroscience*, *16*(4), 702–13.

Northoff, G., Walter, M., Schulte, R. F., Beck, J., Dydak, U., Henning, A., … Boesiger, P. (2007). GABA concentrations in the human anterior cingulate cortex predict negative BOLD responses in fMRI. *Nature Neuroscience*, *10*(12), 1515–1517.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–4.

Ohta, S., Fukui, N., & Sakai, K. L. (2013). Syntactic computation in the human brain: the degree of merger as a key factor. *PLOS ONE*, *8*(2), e56230.

Ojima, S., & Okanoya, K. (2014). The Non-Hierarchical Nature of the Chomsky Hierarchy-Driven Artificial-Grammar Learning. *Biolinguistics*, *8*, 163–180.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, *19*(4), 1730–1737.

Oswa, A., Maeshima, S., Tanemura, J., Sekiguchi, E., & Itakura, T. (2006). Word fluency test in memory clinic ： An efficient tool for diagnosing amnestic

patients. *Higher Brain Function Research*, *26*, 327–333.

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(6), 2522–2527.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291–310.

Phan, K. L., Taylor, S. F., Welsh, R. C., Ho, S. H., Britton, J. C., & Liberzon, I. (2004). Neural correlates of individual ratings of emotional salience: A trial-related fMRI study. *NeuroImage*, *21*(2), 768–780.

Phelps, E. A., Hyder, F., Blamire, A. M., & Shulman, R. G. (1997). FMRI of the prefrontal cortex during overt verbal fluency. *Neuroreport*, *8*(2), 561–565.

Piazza, M., Mechelli, A., Price, C. J., & Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: an fMRI study. *Brain Research*, *1106*(1), 177–88.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*(2), 201–236.

Prins, R., & Bastiaanse, R. (2006). The early history of aphasiology: From the Egyptian surgeons (c.1700 BC) to Broca (1861). *Aphasiology*, *20*(8), 762–791.

Puts, N. A., Edden, R. A., Evans, C. J., McGlone, F., & McGonigle, D. J. (2011). Regionally Specific Human GABA Concentration Correlates with Tactile Discrimination Thresholds. *Journal of Neuroscience*, *31*(46), 16556–16560.

Qin, Y., Carter, C. S., Silk, E. M., Stenger, V. A., Fissell, K., Goode, A., & Anderson, J. R. (2004). The change of the brain activation patterns as children learn algebra equation solving. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(15), 5686–91.

Raichle, M. E. (2013). The Brain's Default Mode Network. *Annual Review of Neuroscience*, *38*(1), 433–447.

Reiser, B. J., Copen, W. A., Ranney, M., Hamid, A., & Kimberg, D. Y. (1998). Cognitive and Motivational Consequences of Tutoring and Discovery Learning. *Cognition and Instruction*, 1–83.

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H. J., ... & Noesselt, T. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology*, 24(21), 2606-2611.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5*(1), 35–53.

Rossor, M., Garrett, N., & Iversen, L. (1980). No Evidence for Lateral Asymmetry of Neurotransmitters in Post-Mortem Human Brain. *Journal of neurochemistry, 35*(3), 743–745.

Santi, A., & Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage*, *51*(4), 1285–1293.

Scheepers, C., & Sturt, P. (2014). Bidirectional syntactic priming across cognitive domains: from arithmetic to language and back. *Quarterly Journal of Experimental Psychology (2006)*, *67*(8), 1643–54.

Scheepers, C., Sturt, P., Martin, C. J., Myachykov, A., Teevan, K., & Viskupova, I. (2011). Structural Priming Across Cognitive Domains From Simple Arithmetic to Relative-Clause Attachment. *Psychological Science*, *22*(10), 1319–1326.

Schenker, N. M., Hopkins, W. D., Spocter, M. A., Garrison, A. R., Stimpson, C. D., Erwin, J. M., … Sherwood, C. C. (2010). Broca's area homologue in chimpanzees (Pan troglodytes): Probabilistic mapping, asymmetry, and comparison to humans. *Cerebral Cortex*, *20*(3), 730–742.

Schmidt, L., Lebreton, M., Cléry-Melin, M. L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biology*, *10*(2), e1001266.

Schneider, E., Maruyama, M., Dehaene, S., & Sigman, M. (2012). Eye gaze reveals a fast, parallel extraction of the syntax of arithmetic formulas. *Cognition*, *125*(3), 475–490.

Segaert, K., Kempen, G., Petersson, K. M., & Hagoort, P. (2013). Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: An fMRI study. *Brain and Language*, *124*(2), 174–183.

Shapiro, K. A., Moo, L. R., & Caramazza, A. (2012). Neural specificity for grammatical operations is revealed by content-independent fMR adaptation. *Frontiers in Psychology*, *3*, 26.

Steedman, M. (2000). *The syntactic process*. Cambridge: MIT press.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language*, *52*(52), 452–473.

Terhune, D. B., Russo, S., Near, J., Stagg, C. J., & Cohen K., R. (2014). GABA predicts time perception. *The Journal of Neuroscience*, *34*(12), 4364–70.

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. a. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience*, *18*(6), 1029–1043.

Tsang, J. M., Dougherty, R. F., Deutsch, G. K., Wandell, B. a, & Ben-Shachar, M. (2009). Frontoparietal white matter diffusion properties predict mental arithmetic skills in children. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22546–51.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–89.

Uylings, H. B. M., Jacobsen, A. M., Zilles, K., & Amunts, K. (2006). Left-Right Asymmetry in Volume and Number of Neurons in Adult Broca ' S Area. *Cortex*, *3*(42), 652–658.

Van de Cavey, J., & Hartsuiker, R. J. (2016). Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition*, *146*, 172–184.

Varley, R. A., Klessinger, N. J. C., Romanowski, C. A. J., & Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(9), 3519–3524.

Wernicke, C. (1874). *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*. Cohn & Weigert, Breslau.

Zhao, Q., Zhou, Z., Xu, H., Chen, S., Xu, F., Fan, W., & Han, L. (2013). Dynamic Neural Network of Insight: A Functional Magnetic Resonance Imaging Study on Solving Chinese "Chengyu" Riddles. *PLOS ONE*, *8*(3), e59351.

Zilles, K., & Amunts, K. (2010). Centenary of Brodmann's map — conception and fate.

Nature Reviews Neuroscience, *11*(2), 139–145.

Acknowledgments