Inhibition processes are dissociable and lateralized in human prefrontal cortex

Lisa Cipolotti[1,2], Barbara Spanò[3], Colm Healy[1], Carina Tudor-Sfetea[1], Edgar Chan[1,] Mark White[4], Francesca Biondo[5], John Duncan[5,6], Tim Shallice[7,8], Marco Bozzali[3]

[1]Department of Neuropsychology, National Hospital for Neurology and Neurosurgery, London, UK

[2]Dipartimento di
Studi di Palermo, Palermo, Italy

[3]Neuroimaging Laboratory, Santa Lucia Foundation, Rome, Italy

[4]Department of Neuroradiology, National Hospital for Neurology and Neurosurgery, London, UK

[5]MRC Cognition and Brain Sciences Unit, Cambridge, UK

[6]Department of Experimental Psychology, University of Oxford, UK

[7]Institute of Cognitive Neuroscience, University College London, UK

[8]International School for Advanced Studies (SISSA-ISAS), Trieste, Italy

Word count: 9244

† Correspondence:

Prof. Lisa Cipolotti

Head of Neuropsychology Department

Box 37

National Hospital for Neurology and Neurosurgery

Queen Square

London WC1N 3BG

Phone 0044 203 4483295

Running Title: Inhibition in Prefrontal Cortex Lesions

Abstract

The prefrontal cortex (PFC) is known to make fundamental contributions to executive functions. However, the precise nature of these contributions is incompletely understood. We focused on a specific executive function, *inhibition*, the ability to suppress a pre-potent response. Functional imaging and animal studies have studied inhibition. However, there are only few lesion studies, typically reporting discrepant findings. For the first time, we conducted cognitive and neuroimaging investigations on patients with focal unilateral PFC lesions across two widely used inhibitory tasks requiring a verbal response: The Hayling Part 2 and Stroop Colour-Word Tests. We systematically explored the relationship between inhibition, fluid intelligence and lesion location using voxel-based lesion symptom mapping (VLSM). We found that PFC patients were significantly impaired compared with healthy comparison group (HC) on both suppression measures of the Hayling and on the Stroop, even when performance on a fluid intelligence test was covaried. No significant relationship was found between patients' performance on each Hayling suppression measure and the Stroop, once fluid intelligence was partialled out, suggesting that the two tests may involve different kinds of inhibition. After accounting for fluid intelligence, we found a significant interaction between tests, Hayling or Stroop, and site, left or right, of PFC damage. This finding suggesting lateralized functional organization was complemented and extended by our VLSM results. We found that performance on both Hayling suppression measures significantly relied on the integrity of a similar and relatively circumscribed region within the right lateral PFC, in the right lateral superior and middle frontal gyri. In stark contrast, performance on the Stroop relies on the integrity of left lateral superior and middle frontal gyri. Thus, lesion location, right or left PFC, is critical in producing impairments on two inhibitory tasks loading similarly on verbal control. This suggests that the two suppression measures of the Hayling and the Stroop are likely to assess dissociable components of executive functions, related to anatomically defined and lateralized PFC circuits. Our findings also suggest that inhibition may actually comprise qualitatively different forms with different neural substrates. This has clinical implications for the diagnosis and treatment of disinhibition impairments, a common behavioural problem caused by PFC lesions. Our results highlight the need to assess inhibition using a variety of tasks and to develop different types of treatments.

Highlights:

- Right or left PFC are critical for performance on the Hayling and Stroop tasks respectively
- The Hayling and the Stroop tasks are likely to assess dissociable components of executive functions, related to anatomically defined and lateralized PFC circuits
- Inhibition may comprise qualitatively different forms with different neural substrates

Key words: Executive function, Inhibition, Fluid Intelligence, Prefrontal Cortex, Hayling and Stroop

Abbreviations: ACC = anterior cingulate cortex, BA = Brodmann Area, CVA = cerebrovascular accident, GNT = Graded Naming Test, HC = healthy comparisons, IQ = Intelligence Quotient, LF = left frontal, MNI = Montreal Neurological Institute, NART = National Adult Reading Test, No = Number, PFC = prefrontal cortex, RAPM = Raven's Advanced Progressive Matrices, RF = right frontal, RIFG = right inferior frontal gyrus, ROI = region of interest, RT = reaction time, SD = standard deviation, TMT-A = Trail Making Test, Part A, VSLM = voxel-based lesion mapping, WAIS-R = Wechsler Adult Intelligence Scale-Revised

1. Introduction

Executive functions refer to a variety of general purpose control mechanisms thought to modulate and organize more basic cognitive sub-processes to achieve effective behaviour (e.g. Stuss & Levine, 2002). The prefrontal cortex (PFC) is widely acknowledged to make a fundamental contribution to executive functions. However, the precise nature of this contribution is incompletely understood (e.g. Hornberger & Bertoux, 2015). It is likely that a prefrontally located executive system critical for non-routine behavior may be divided into anatomically separable subsystems (e.g. Stuss and Alexander, 2007; Shallice et al., 2008). Models have been proposed suggesting a functional organization of the PFC along the dorsal-ventral (Petrides, 2005) or rostro-caudal (Badre, 2008; Badre & D'Esposito, 2009) axis of the PFC. Lesion data also suggested that the functional organization of the PFC may be lateralized (Stuss & Alexander, 2007; Shallice & Gillingham, 2012). One prime example of this is that executive functions linked to verbal or non-verbal generation have been associated with left or right PFC respectively (e.g. Robinson et al., 2012).

Despite the likelihood of some anatomical separation, some theories suggest that the PFC carries out general control processes to match the requirements of the task being undertaken, independently of the type of information being processed (e.g. Duncan, 2001; Miller & Cohen, 2001). A large PFC-parietal network, named the *multiple-demand network*, has been shown to be associated with a wide range of cognitive operations in functional imaging work. This putative network has been proposed to be the seat of general fluid intelligence (*g*; e.g. Woolgar et al., 2010). It is well known that fluid intelligence is positively correlated with tests of executive functions and is impaired following frontal lesions (Duncan et al., 1995). However, very few studies that assume executive functions are separable and associated with different sub-regions of the PFC have actually investigated whether executive impairments in frontal patients can be explained by a fluid intelligence loss. A notable exception is represented by the study of Roca and colleagues (2010). The authors reported that, for several executive tests, such as the Wisconsin Card Sorting Test, Verbal fluency and the Iowa Gambling Task, when performance on a fluid intelligence test was taken into account, no significant difference was found in the performance of frontal patients and HC. Therefore, it remains important to establish the extent to which a loss of fluid intelligence can account for executive impairments in frontal patients.

In this paper we are concerned with a specific executive process, *inhibition*, generally referred to as the ability to suppress inappropriate responses, a definition that we will adopt throughout our paper. *Inhibition* is widely accepted as one of the key components of executive functions (e.g. Miyake et al., 2000). Patients with PFC lesions may present impairments in tasks requiring inhibitory control and may manifest problems resulting from deficits in inhibition, such as inappropriate and/or perseverative behaviour. Functional imaging and animal studies have long been used to investigate inhibition (e.g. Aron et al., 2004, 2014). However, there is a paucity of lesion studies on inhibition. So far, none has systematically investigated the relationship between fluid intelligence, inhibitory tasks and lesion location, despite the fact that some of the inhibitory tasks used are known to be related to fluid intelligence. Moreover, these few studies have used different tasks varying on important dimensions such as the response modality and have reported inconsistent results. Therefore, it remains unclear which PFC areas contribute to inhibition and how.

For example, Miyake and colleagues (2000) suggested that *inhibition* is related to performance on the Tower of Hanoi test, known to be significantly related to fluid intelligence (e.g. Zook et al., 2004). Aron and colleagues (2003) used a stop-signal task to investigate inhibition in patients with right PFC damage and HC and reported a significant correlation between right inferior frontal gyrus (RIFG) lesions and stop-signal reaction time (RT). They suggested that the RIFG was critical for inhibitory control in general. Unfortunately, it remains unclear whether the RIFG does indeed harbour an inhibitory system (e.g. see Parvizi, 2012, for a contrasting view). In a subsequent study, using essentially the same set of patients as previously reported, Aron and colleagues (2004) adopted a test of task-set switching. They found that right frontal patients made more errors than left frontal patients, at a short response stimulus interval, supporting the idea that inhibition was mediated by the RIFG. However, Picton and colleagues (2007), using a go-no-go task, found that the incidence of false alarms was nearly three times higher, and significantly so, in patients with lesions in left superior medial frontal cortex than in patients with right inferior frontal lesions, who in turn were not very different from HC. A medial frontal involvement also seemed to be critical in a much earlier study (Drewe, 1975). These findings raise the possibility that other areas of the PFC may be involved in inhibition.

The Stroop is one of the most widely used tests in neuropsychology and it is considered a classic ‘inhibitory’ task (Stroop, 1935). The Stroop Colour-Word part of

this test requires participants to name the ink colour in which a colour name is printed when the written name is incongruent (e.g. the word 'yellow' printed in red and the participant has to name the colour instead of reading the word). Thus, successful performance has long been thought to measure the ability to inhibit pre-potent verbal responses (e.g. Cohen & Servan-Schreiber, 1992; Logan, 1994; Friedman & Miyake, 2004). It has also been suggested that this task requires other abilities such as conflict monitoring (e.g. Botvinick et al., 2001; De Pisapia & Braver, 2006), modulation of strategic control (Kerns et al., 2004), working memory (e.g. Kimberg & Farah, 1993) and general goal maintenance (e.g. Cohen & Servan-Schreiber, 1992; West & Baylis, 1998). Similarly to the Tower of Hanoi Test, the Stroop has also been linked to fluid intelligence. It has been reported that in HC, performance on the Stroop is associated with performance on the Wechsler Adult Intelligence Scales-Revised (WAIS-R), suggesting that both are measures of *g* (e.g. Obonsawin et al., 2002).

For a long time, evidence for functional specialization within the PFC was conflicting for Stroop. In the first neuropsychological study of Stroop, which examined performance in 118 patients with focal lesions, Perret (1974) reported an impairment in patients with lesions in left dorsolateral PFC. Stuss and colleagues (2001) also reported poor Stroop performance after left dorsolateral and superior medial lesions, particularly involving the right supplementary motor area. In contrast, Vendrell and colleagues (1995) reported impairments on the Stroop in patients with lesions in right lateral PFC. The latter two studies included a sizeable number of patients with traumatic brain injury, patients with bilateral lesions, and patients with lesions extending beyond the PFC, raising the possibility that the findings may partially reflect damage beyond the identified area. Several small patient series with more or less selective lesions of anterior cingulate cortex (ACC) have also reported inconsistent results (e.g. Swick & Turken, 2002; Fellows & Farah, 2005; Baird et al., 2006). A number of recent studies have used Voxel-based Lesion Symptom Mapping (VLSM) to investigate the structure-function relationship in the Stroop. This is a method of image analysis allowing an operator-independent measurement of association between anatomical localization of brain tissue damage and patients' performance in a specific cognitive task (Bates et al., 2003). These studies have suggested that the left PFC is involved in the Stroop Colour-Word test. Thus, Tsuchida and Fellows (2013) investigated the performance of patients with PFC lesions on several executive tasks, including the Stroop and found that the left lateral PFC substantially contributed to performance on the Stroop (for additional findings

suggesting a left PFC involvement, see Demakis, 2004; Derrfuss et al., 2005). In a subsequent study, Geddes and colleagues (2014) documented impairment in the Stroop task on a small number of patients with left ventral lateral PFC damage but not in patients with right ventral lateral PFC damage. Interestingly, the authors reported that patients with right ventral lateral PFC damage were impaired on the Eriksen Flanker tests (Eriksen & Eriksen, 1974), also requiring suppression of pre-potent responses. A further study using VLSM in a large sample of patients with PFC lesions (n=165) reported that the left dorsolateral frontal cortex was associated with poor performance on the Stroop (Glascher et al., 2012). Altogether, these findings have been taken to challenge the notion that the right lateral PFC is critical for inhibitory control in general (Tsuchida & Fellows, 2013) and to support a role of inhibitory control to the left dorsolateral PFC (Glascher et al., 2012).

Part 2 of the Hayling Sentence Completion Test has also been considered to assess response inhibition/suppression, as well as strategy generation (Burgess & Shallice, 1996). It requires the patient to complete sentences by providing words that are unrelated to the sentence frame (e.g. '*London is a very busy*…' could be completed by saying… '*banana*'). Frontal patients may produce Suppression errors (e.g. '*London is a very busy*…' may be completed with '…*city*…') and may require longer reaction time (RT2). These deficits are thought to demonstrate a difficulty in inhibiting the pre-potent natural-completion word. Roca and colleagues (2010) reported that, in a shorter version of the Hayling, when fluid intelligence was partialled out, frontal deficits remained. The authors suggested that the unique deficit captured by the Hayling may be associated with anterior, especially right, frontal lesions. Volle and colleagues (2012) searched for regions most associated with Suppression errors and suggested a focus in right Brodmann Area (BA) 11. Hornberger and colleagues (2011) reported impaired Suppression error scores in frontotemporal dementia patients, which correlated with atrophy in ventromedial orbitofrontal cortex, subgenual, as well as anterior temporal and medial frontal grey matter. We recently documented a significantly greater number of Suppression errors and elevated RT2 in patients with right inferior and/or middle frontal gyri lesions when compared to HC (Robinson et al., 2015). In a subsequent study, we compared the performance of a small sample of PFC patients with right lateral or orbitofrontal lesions. We replicated the Suppression errors and elevated RT2 for the right lateral patients. In contrast, we found that orbitofrontal patients were unimpaired on both measures (Cipolotti et al., 2015a).

Notably, some of the discrepancies reported above may be due to the fact that inhibition has been tested with tasks requiring motor or verbal responses. Hemispheric differences could in part depend on response modality. The neuroimaging techniques used so far also differ greatly and may, in part, account for the reported disparities (e.g. ROI; VLSM; MRI/CT scan qualitative analysis). For example, Volle and colleagues (2012), using two different imaging techniques, reported somewhat different PFC areas involved with Part 2 of the Hayling. Using AnaCOM, a voxel-by-voxel lesion mapping method with some similarities to VLSM, the authors found that deficits in Part 2 of the Hayling were associated with lesions in the RIFG and in BA11. However, using VLSM they only found a small cluster in right BA11. Moreover, only one type of inhibitory task was administered to the frontal patients. This made an investigation of potential differences in performance in the same sample of frontal patients according to the type of inhibitory task used impossible. Moreover, it did not allow for the assessment of relations between these potential differences, their associated lesion location and fluid intelligence. Unsurprisingly, it remains unclear if inhibition can be considered a functionally separable executive function associated with a specific PFC region.

The aim of our study was to investigate the contribution of the PFC to inhibition. For the first time we evaluated whether fluid intelligence contributed to the performance of frontal patients on two clinically and experimentally widely used inhibitory tasks, namely the Stroop Colour-Word Test (called the Stroop in this paper) and Part 2 of the Hayling Sentence Completion Test (called the Hayling in this paper). Moreover, for the first time we attempted to examine the pattern of performance across these two types of inhibitory tasks requiring a verbal response in a sample of PFC patients, using the same neuroimaging investigations. This approach allowed us to evaluate whether distinct PFC regions were differentially and critically involved in performance on these two tasks. If the PFC regions collectively contributed to a shared underlying mechanism critical for performance on the Stroop and the Hayling, the expectation was to find a common pattern of lesion-symptom association across both tasks. Alternatively, if different PFC regions made distinct contributions to these two tasks, we expected to be able to dissociate performance on the Stroop and Hayling task according to lesion location.

We retrospectively examined performance on the Stroop and the Hayling tests in a sample of patients with a unilateral PFC lesion. Our cognitive investigations examined the role of fluid intelligence and compared the performance of left and right frontal

patients. Our neuroimaging investigations used VLSM as well as region of interest (ROI) group comparisons analyses.

2. Materials and methods

2.1. Participants

For the **Cognitive and VLSM Investigations**, 164 patients with unilateral, focal lesions confined to the frontal lobes, resulting from a cerebrovascular accident (CVA) or a brain tumour, who attended the Neuropsychology Department at the National Hospital for Neurology and Neurosurgery, Queen Square, London, were screened for eligibility. The following exclusion criteria were employed both for the Cognitive and the VLSM investigations: i) age at the time of cognitive testing > 80 years, due to the availability of age-matched HC data and standardised age norms only for patients up to 80 years, ii) current or previous psychiatric disorders, iii) previous neurological disorders including CVAs or tumours, iv) presence of metastatic tumours, v) chemotherapy previous to the tumour resection as well as following resection, prior to the cognitive investigations, vi) gross visual (i.e., cortical blindness), perceptual (i.e., neglect; agnosia), motor (i.e., hemiplegia) or language (i.e., dysphasia) impairment, vii) previous head trauma, viii) history of alcohol or drug abuse, ix) no MRI or CT scan results available, x) a score below the 5^{th} percentile on a test of fluid intelligence (Raven's Advanced Progressive Matrices, RAPM, Raven, 1976), xi) non-native English speakers were only included in the study if they obtained a score at or above the 25^{th} percentile on the National Adult Reading Test (NART, Nelson, 1982), to ensure that their English abilities were sufficient to cope with task demands. In addition, for the **Cognitive Investigations**, only patients with both Stroop and Hayling Sentence Completion Tests data available were included. These patients did not have high-resolution MRI, and accordingly were not used for VLSM. For the **VLSM Investigations**, we included patients with either Stroop or Hayling data (no patients having both), along with high-resolution T1-weighted MRI.

For the **Cognitive Investigations**, application of the exclusion criteria resulted in 30 frontal patients (15 left and 15 right) with Stroop and Hayling Sentence Completion Tests (see Table 1). This is an entirely new sample of frontal patients never reported before. As far as we are aware this is the first time that data on both Stroop and Hayling tests have been documented in the same sample of patients. For the **VLSM**

Investigations, application of the exclusion criteria resulted in 31 patients (14 left and 17 right frontal patients) with Stroop and a second group of 27 patients (13 left and 14 right frontal patients) with Hayling. Their demographic and cognitive characteristics were similar to those reported for the patients used for the Cognitive Investigations (see Table 1). Some aspects of the cognitive profile of 7 out of the 31 patients with Stroop only data available have been previously reported in the context of a larger study (n=100 frontal patients) investigating the impact of different aetiologies on the cognitive performance of frontal patients (Cipolotti et al., 2015b). Some aspects of the cognitive profile of 8 out of the 27 patients with Hayling only data available have also been previously reported in the context of a smaller study (n=11 frontal patients) contrasting the performance of right lateral and orbitofrontal patients on the Hayling test (Cipolotti et al., 2015a). Notably in these two previous studies different neuroimaging techniques were used (MRI/CT qualitative analysis and ROI). This is the first time we have conducted a VLSM analysis. For all patients, the diagnosis was confirmed by neurological investigation.

The aetiologies of the frontal lesions were stroke (n=10 for the cognitive investigations; n=8 for the VLSM/Stroop and n=1 for the VLSM/Hayling), high-grade tumour (n=3 for the cognitive investigations; n=7 for the VLSM/ Stroop and n=6 for the VLSM/Hayling), low-grade tumour (n=11 for the cognitive investigations; n=10 for the VLSM/Stroop and n=6 for the VLSM/Hayling) and meningioma (n=6 for the cognitive investigations; n=6 for the VLSM/Stroop and n=14 for the VLSM/Hayling). All tumour patients had undergone tumour resection prior to neuropsychological testing. Importantly for the current study, we have previously documented no significant differences in the performance of frontal patients with CVA, high- or low-grade tumour and meningioma on the RAPM, Stroop Colour-Word and Graded Naming Tests (GNT). This suggests that grouping together of frontal patients with different aetiologies is methodologically justifiable (Cipolotti et al., 2015b).

Data from 60 HC who did not significantly differ from the frontal patients in terms of age, gender, NART IQ and years of education was also reviewed. The HC were recruited from the relatives of patients attending the Neuropsychological Department and their data were also retrospective in nature but contemporary to the frontal patients' data.

The HC were not given monetary compensation for participation. Retrospective recruitment of patients was approved by the National Hospital for Neurology and Neurosurgery and the Institute of Neurology Joint Research Ethics Committee, University College London Hospitals NHS Trust Research and Development Directorate.

Insert Table 1 about here

2.2. Cognitive Investigations

We retrospectively reviewed the cognitive performance of the 30 frontal patients with Stroop and Hayling data available and the 60 HC on a single neuropsychological assessment comprising well-known tests with published standardised normative data. All tests were administered in the published standard manner. Data were available on the following baseline tests: NART, estimating optimal pre-morbid functioning; the RAPM test assessing fluid intelligence (No. of correct answers out of 12 items; Arthur and Day, 1994); the GNT assessing nominal functions (No. of correct answers; McKenna and Warrington, 1980), Part A of the Trail-Making Test assessing speed of information processing (TMT-A; Reitan, 1992); and the Phonemic Fluency test (No. of words named starting with the letter S in 60 seconds; Benton, 1968). In addition, Section 2 of the Hayling Sentence Completion Test and the incongruent condition of the Stroop had also been administered.

Following Burgess and Shallice (1996), we calculated two scores for section 2 of the Hayling: Suppression errors (Category A Errors or regular sentence completions = 3 points, Category B Errors or related completions = 1 point) and total Suppression RT2. Scaled scores ranged from 1 to 10 with the points corresponding with the following percentiles: 1 = *out of normal range* or <1st percentile; 2 = 1st percentile; 3 = 5th percentile; 4 = 10th percentile; 5 = 25th percentile; 6 = 50th percentile; 7 = 75th percentile; 8 = 90th percentile; 9 = 95th percentile; 10 = 99th percentile.

We used the Trenerry et al. (1989) version of the Stroop test which consisted of 112 colour words (red, green, blue or tan), each printed in one of the three incongruent ink colours (i.e. no word is printed in its matching colour). The words were arranged in four equal columns on one A4 sheet. We recorded the total number of ink colours correctly named in two minutes. If participants correctly named all ink colours in less than two minutes, their score was prorated to reflect the number of colours they would have achieved in two minutes. This prorated score was used as a dependent variable for patients and HC. On the basis of normative data, participants' Stroop performance was classified as impaired if scores were below the 5th percentile (Trenerry et al., 1989).

For all HC, data were available on the NART, RAPM and the GNT. For 38 HC, data were available on the TMT-A and for 20 HC, on the phonemic fluency test. For 21 HC, data were available for the Stroop, but not the Hayling. For 39 HC, data were available on the Hayling, but not the Stroop.

For the 30 patients included in the Cognitive Investigations, two neurologists (MB and BS) who were blind to the experimental results reviewed hard copies or digital records of the MRI or CT brain scans. For 26 out of 30 patients, either MRI T1-weighted (n=20) or CT scans (n=6) were available. For 4 out of 30 patients, lesion lateralization was based on the clinical neuroradiological report. MRI images were obtained by either 1.5T GE Medical System (n=16) or 3T Siemens (n=4) magnetic resonance scanners. CT images were obtained by Philips Medical System (n=2), GE Medical System (n=2) or Siemens (n=2) spiral CT systems. The assessment of our patients' frontal lesions was based on detailed anatomical localization using standard atlases (Duvernoy, 1991). All lesions were entirely located within the frontal lobe. Each frontal patient was coded for the presence of lesion (if at least 25% of an area was affected) and oedema in each hemisphere in the anterior and posterior portion of nine Left and Right frontal regions (18 areas in total; for further details see Turner et al., 2007). For present purposes, the nine left and right brain regions were collapsed together and in the lateralization analysis we divided the patients into Left (LF; n = 15) and Right (RF; n=15) Frontal groups.

2.3. VLSM Investigations

We used VLSM to explore whether separable PFC regions made distinct contributions to performance on the two suppression measures of the Hayling and on the Stroop. Unfortunately, we did not have a sample of patients with Hayling and Stroop data and digital MRI scans allowing for a VLSM analysis. As mentioned above, to be eligible for VLSM analysis, patients had to have a T1-weighted MRI scan, and: 1) Hayling Suppression errors and Hayling Suppression RT2 or 2) Stroop. In this subsample of patients, MRI images were obtained using MR scanners either operating at 1.5T (n=23 for Hayling and n=31 for Stroop) or 3T (n=4 for Hayling and n=0 for Stroop).

For all patients, focal lesions were first identified and outlined (using a semi-automated local threshold contouring software; Jim 5.0, Xinapse System, Leicester, UK, http://www.xinapse.com) by an expert observer (BS), who was blinded to the patients' performance. For every patient, a lesion mask was created (by assigning a value of 1 to

every voxel corresponding to a lesion and 0 elsewhere) and normalized to the Montreal Neurological Institute (MNI) space, as previously described (Torso et al., 2015). A lesion map, indicating the number of patients with a lesion in a given area, was obtained by combining every patient's lesion masks (see Figure 1).

Insert Figure 1 about here

Next, three separate VLSM analyses were run using the VLSM2 Matlab toolbox (version 2.55, http://neuroling.arizona.edu/resources.html) to find significant associations of lesioned voxels and Suppression errors, Suppression RT2 and Stroop scores. Two of the VLSM analyses included those patients (n=27) with Suppression errors and RT2 data available. The third one included those patients (n=31) with Stroop data available. In these analyses, the neuropsychological variable taken into consideration (i.e., Hayling Suppression errors, Hayling Suppression RT2 and Stroop) was used as the variable of interest and age was always entered as a covariate of no interest (Bates et al., 2003).

Because of the low probability of patients having damage in the same voxel, statistical power in VLSM analyses is often low. Thus, in line with other studies, we accepted a low power in order to be able to include a larger area of the brain. For our study we chose the minimum number of patients with overlapping lesions to be 4 at any voxel (see Knutson et al., 2013; Glascher et al., 2009 who used a similar criterion). The resulting *t*-statistic maps were corrected for multiple comparisons at cluster level using permutation tests (Kimberg et al., 2007) with 1000 repetitions, accepting as significant p values of less than 0.05. To ensure that there was sufficient left-right symmetry of power, we created statistical power maps using VLSM 2 software. The power map reflects the probability to detect a true relationship between damage in a certain region and a behavioural score of interest. The power varies as a function of the proportion of patients presenting with damage in that region (see Kimberg et al., 2007 for further discussion). Examination of the resulting power maps revealed relative left-right symmetry (see Figure 2).

Insert Figure 2 about here

2.4. Region-of-interest Investigations

The VLSM analyses were supplemented with two more conventional group comparisons for the same group of PFC patients. These two additional post-hoc analyses were conducted to compare our data with previously published studies. In particular, to address the issue as to whether damage to the RIFG is critical for inhibitory control in general we compared the performance on the Stroop between patients with lesions including (n=8) or not including (n=23) the RIFG using ANOVA (Aron et al., 2003, 2004). Similarly, to address the issue of whether ACC plays a role in the performance on the Stroop we compared the performance of patients with lesions including (n=16) or not including (n=15) the ACC (e.g. Botvinick et al., 2001; Carter et al., 2000). The anatomical ROIs for the RIFG and ACC were defined in standard-space using the anatomical Atlas included in FSL library (fsl.fmrib.ox.ac.uk/fsl). The RIFG's size was
24.3 mL and its center of gravity in mm standard-space coordinates was: 49.5; 21.8; 12.9. The ACC's size was 22.7 mL and its center of gravity in mm standard-space coordinates was: 0.7; 19.9; 24.2.

3. Results

3.1. Cognitive Investigations

3.1.1. Demographic and General Cognitive results

We conducted independent samples t-test or chi-square analyses to examine whether frontal patients and HC differed on demographic and general cognitive variables. Frontal patients and HC were well-matched for age ($t(88)=-0.292$, $p=.771$), gender (χ^2 (n=90, df=1)=0.584, $p=.455$), NART IQ ($t(45.02)=1.605$, $p=.116$) and years of education ($t(88)=0.219$, $p=.827$). The mean time between damage and assessment was 8.87 months (standard deviation (SD) =20.85 months). We found significant differences on performance on the RAPM ($t(82)=3.194$, $p=.002$) and TMT-A ($t(60)=-2.022$, $p=.048$). There was no significant difference between frontal patients and HC on the GNT ($t(43.01)=1.364$, $p=.180$) or on Fluency S ($t(46)=1.045$, $p=.301$; see Table 1). Crucially, there was also no significant difference between left and right frontal patients on any demographic variables or performance on general cognitive tasks.

3.1.2. Part 2 of the Hayling and Stroop: Overall results- PFC patients and HC

As stated above, for the Hayling we calculated Suppression errors and Suppression RT2 as detailed by Burgess and Shallice (1996). For the Stroop, the prorated number of colours correctly named in the conflict condition was used. Frontal patients were impaired compared to HC on Suppression errors (t(48.99)=2.022, p=.049), Suppression RT2 (t(38.77)=3.621, p=.001) and the Stroop (t(40.89)=2.900, p=.006; see Table 2). We ran an ANCOVA with RAPM as a covariate to examine the potential contribution of fluid intelligence to these differences. We found that RAPM was a significant covariate, separately for each measure (p<0.01). Nevertheless a significant difference remained in Suppression RT2 and Stroop performance between frontal patients and HC, even when we covaried for performance on the RAPM (F(1,42)=7.393, p=.009 and F(1,58)=13.115, p=.001 respectively).

Qualitatively, we found that 30% of frontal patients (n=9) obtained a score below the 5th %ile on Suppression errors**,** 20% (n=6) on Suppression RT2 and 20% (n=6) on the Stroop. In contrast, no HC obtained scores below the 5th%ile on any of these three measures. Of note, four patients with an impaired performance (<5th percentile) on Suppression errors and Suppression RT2 performed well (>70th percentile) on the Stroop. Conversely, three of the patients with an impaired performance on the Stroop (<5th percentile) performed well on Suppression errors and Suppression RT2 (>50th percentile). This suggests that impairments on the suppression measures of the Hayling and on the Stroop may be doubly dissociated.

We also investigated if there were significant differences in the performance between right and left frontal patients and healthy controls on the Hayling and Stroop tests. We used univariate ANOVAs with post-hoc Bonferroni corrections. We found that, on Suppression errors and Suppression RT2, right frontal patients were significantly impaired compared to HC. However, left frontal patients were not significantly impaired compared to HC (Suppression errors: F(2,48)=3.146, p=.05; right frontal patients vs. HCs p=.049; left frontal patients vs. HCs p=1.000; Suppression RT2 (F(2,48)=5.878, p=.005; right frontal patients vs. HCs p=.004; left frontal patients vs. HCs p=.165). In contrast, on the Stroop, we found that left frontal patients were significantly impaired compared to HC. However, right frontal patients were not significantly impaired compared to HC (F(2,63)=5.316, p=.007; left frontal patients vs. HCs p=.010; right

frontal patients vs. HCs p=.201). When we entered RAPM as a covariate, we found that RAPM contributed significantly to the performance in both the Hayling and Stroop, as before (p<.05). Importantly, after covarying for RAPM, a significant difference remained in Suppression RT2 between right frontal patients and HC (F(2,41)=6.461, p=.004; right frontal patients vs. HCs p=.003) and in Stroop between left frontal patients and HC (F(2,57)=6.867, p=.002; left frontal patients vs. HCs p=.003).

Insert Table 2 about here

3.1.3. Part 2 of the Hayling, Stroop and Fluid intelligence (RAPM): correlation results in PFC patients

Given the inhibitory aspect thought to be involved in Section 2 of the Hayling and on the Stroop, we first ran two-tailed Pearson's correlations to investigate if, among frontal patients, there was a relationship between performance on Suppression errors, Suppression RT2, and Stroop. We found significant correlations between: Suppression errors and Suppression RT2; Suppression errors and Stroop; Suppression RT2 and Stroop (see Table 3). Secondly, we investigated if, in frontal patients, there was a relationship between the performance on a test of fluid intelligence test (RAPM) and performance on Suppression errors, Suppression RT2, and Stroop. As expected, significant correlations were found between Suppression errors and RAPM (r=.445, p=.029), Suppression RT2 and RAPM (r =.522, p=.009), and Stroop and RAPM (r=.549, p=.008), demonstrating that better performance on Hayling Part 2 and Stroop was associated with better performance on the RAPM. We then ran a partial correlation to assess whether our three measures would still be correlated, once the variance explained by fluid intelligence was partialled out. We found that the relationship between Suppression errors and Suppression RT2 remained significant. However, the relationships between Suppression errors and Stroop, and between Suppression RT2 and Stroop, were no longer significant (see Table 3).

These results suggest that fluid intelligence mediates the relationship between the two measures of the Hayling section 2 and Stroop. However, once fluid intelligence has been taken out, no significant relationship remains between performances on our two different tests.

Insert Table 3 about here

3.1.4. Comparing Left and Right frontal patients on Suppression errors, Suppression RT2, and Stroop, after accounting for fluid intelligence (RAPM).

We investigated if, after accounting for fluid intelligence (RAPM), there was a difference in the performance of left and right frontal patients on the three inhibition measures: Suppression errors, Suppression RT2, and Stroop. Notably in our sample of PFC patients we found no significant difference between the performance of the right and left frontal patients on the RAPM ($t(22)=1.434$, $p=.166$). To allow us to compare our three measures, we converted the patients' scores into z-scores, based on the performance on the HC. We then conducted three separate 2x2 mixed-method ANOVAs with type of measure as the within-groups factor, site of damage as the between groups factor, and fluid intelligence (RAPM) as the covariate. The within-groups comparison compared Suppression errors with Stroop, Suppression RT2 with Stroop and Suppression RT2 with Suppression errors. The between-group factor compared left and right frontal lobe lesions. We investigated the main effects and interactions between these two factors for each of the three ANOVAs. Alpha-levels were set as $p<.05$.

Suppression errors and Stroop: There was no main effect for Type of measure used ($F(1,19)=3.356$, $p=.083$) or Site of Damage ($F(1,19)=0.435$, $p=.518$). However, we found a significant interaction between the two factors ($F(1,19)=8.796$, $p=.008$). Inspection of the mean z-scores revealed that left frontal patients performed worse on the Stroop compared to Hayling Suppression errors. The opposite pattern of results was obtained for the right frontal patients. Right frontal patients performed worse on the Hayling Suppression errors and better on the Stroop (see Figure 3a). Post-hoc comparison with Bonferroni adjustments showed that the difference in performance between the two tasks was statistically significant for left frontal patients ($p=.002$) but not the right frontal patients ($p>0.1$).

Suppression RT2 and Stroop: Again we found no main effect for Type of measure used ($F(1,19)=0.490$, $p=.493$) or Site of damage ($F(1,19)=0.110$, $p=.307$). However, we found a significant interaction between the two factors ($F(1,19)=5.654$, $p=.028$). Inspection of the z-scores revealed a pattern of results almost identical to the one

reported above (see Figure 3b). Post-hoc comparison with Bonferroni adjustments showed that the difference in performance between the two tasks was statistically significant for right frontal patients (p=.007) but not left frontal patients (p>.10).

Suppression Error and Suppression RT2: We found a significant main effect for Type of Measure (F(1,21)=5.844, p=.025) with performance on Suppression RT2 being worse than Suppression errors. However, we found no main effect for Site of damage (F(1,21)=2.898, p=.103) and no significant interaction between these two factors (F(1,21)=0.547, p=.468).

These results suggest that, after accounting for fluid intelligence, left and right frontal lesions make distinct contributions to the executive processes involved in the Hayling and Stroop.

Insert Figure 3 about here

3.2 VLSM Investigations

As a preliminary to our VLSM investigations, we assessed whether the size of the lesion differed between left and right frontal patients. We ran t-tests on our two patient samples. We found no significant differences in lesion size between left and right frontal patients on either sample (Hayling t(25)=-1.733, p=.112; Stroop t(29)=1.683, p=.107). Given these results, it seems unlikely that lesion size contributed to any VLSM lateralisation effects. Thus, lesion size was not entered as a covariate in our VLSM analyses to improve power as suggested by Kimberg and colleagues (2007). We also investigated whether size of lesion differed by gender and laterality, given previous research (Piefke et al., 2005; Stevens & Hamann, 2012; Sutterer et al., 2015; Tranel et al., 2005). A two-way ANOVA with gender and laterality as factors and lesion size as the dependent variable revealed no main effect or interaction for either the Hayling sample (p>0.1) or the Stroop sample (p>0.1). As such, gender was not considered in the main VLSM analyses.

Figure 4 shows the main VLSM results we obtained separately for Hayling Suppression errors, Hayling Suppression RT2 and Stroop.

VLSM Hayling Suppression errors: A significant association was found between Suppression errors and the presence of lesion in the right frontal lobe. In particular, we

found a significant association with the right superior and middle frontal gyri (BA 6, 8, 9, and 46), including both gray and white matter (see Figure 4A).

VLSM Hayling Suppression RT2: In a similar fashion to the result obtained for the Suppression errors we found a significant association between patients' performance on RT2 and the presence of lesion in the right frontal lobe. In particular lesion distribution involved the right superior and middle frontal gyri (BA 6, 8, 9), including both gray and white matter (see Figure 4B).

VLSM Stroop: We found a significant association between poor performance on Stroop and the presence of lesion in the left frontal lobe. In particular lesion distribution involved the left superior frontal gyrus (BA 6, 8, 9), including both gray and white matter (see Figure 4C).

Insert Figure 4 about here

3.3. Region-of-interest Investigations

Using ANOVA, we examined the performance on the Stroop of PFC patients with lesions involving the RIFG and compared it with that of PFC patients whose lesions did not include the RIFG. We found no significant difference between the performance of patients with lesions involving or not involving the RIFG ($F(1,29)=1.823$, $p=.187$). Similarly, we examined the performance on the Stroop of PFC patients with lesions involving the ACC and compared it with that of PFC patients whose lesions did not involve the ACC. Again we found no significant difference in the performance of these two groups of patients ($F(1,29)=1.356$, $p=.254$). Of note, none of the patients with lesions involving the RIFG scored below the 5^{th} percentile on the Stroop.

4. Discussion

To our knowledge, this is the first lesion study reporting a systematic investigation of the performance of focal unilateral frontal patients on two widely used inhibitory tasks: the Hayling Part 2 (Suppression errors and Suppression RT2) and Stroop. Our study systematically explored whether fluid intelligence was a substantial contributor to impairment on both tasks. As far as we are aware this analysis has not been attempted before for both tasks. Moreover, our combination of cognitive, VLSM and ROI analyses allowed us to investigate the locations within the PFC that are critical to

performance on these two tasks. Our results further our understanding of the functional organization of the PFC and provide some novel insight into frontal lobe impairments and their associated regions of damage.

The main findings of our Cognitive Investigations were that our PFC patients were significantly impaired compared with HC on both suppression measures of the Hayling and on the Stroop. Investigation of the relationship between performance on a fluid intelligence test (RAPM), both suppression measures of the Hayling and the Stroop revealed that there were significant correlations between performance on the RAPM and on our two inhibitory tasks. However, the difference in performance between PFC patients and HC on our two inhibitory tasks remained significant, even when RAPM performance was covaried. Roca and colleagues (2010) have previously reported that fluid intelligence is a substantial contributor to PFC patients' executive impairments. Only for a handful of frontal tasks did the authors report impairments in PFC patients, after partialling out the contribution of fluid intelligence. One such task was a shortened version of the Hayling test (3 sentences) which was administered to a much smaller sample of frontal patients (n=15) than the one investigated in our current study. Our results add to the findings of Roca and colleagues, suggesting that frontal patients' impairment on the Suppression errors and Suppression RT2 of the Hayling cannot be fully explained by fluid intelligence. Moreover, our findings also demonstrate that the Stroop is another 'executive' task for which impairment cannot be accounted for entirely by fluid intelligence. With regards to the relationship between the three variables of interest, once fluid intelligence was partialled out the correlation between the two Hayling measures remained significant, as would be expected if they have a common component, separate from fluid intelligence. In contrast, we found no significant relationship between patients' performance on each Hayling measure and the Stroop, once fluid intelligence was partialled out. This finding suggests that our two inhibitory tasks are unlikely to rely solely on a common component, such as fluid intelligence, or a general inhibition mechanism.

Interestingly, we found evidence for lateralized functional organization both in the Cognitive and in the VLSM investigations. In our Cognitive Investigations, we found that right frontal patients were significantly impaired compared with HC on both suppression measures of the Hayling whilst left frontal patients were significantly impaired compared with HC on the Stroop. Moreover, after accounting for fluid intelligence, there was a significant interaction between performance on the two

suppression measures of the Hayling and Stroop and the site of PFC damage (left or right). These results further suggest that the type of inhibitory task used and the side of frontal lesion result in differences in frontal patients' performance. In a handful of patients we qualitatively observed a double dissociation between impaired performance on Suppression errors and Suppression RT2 on the one hand and preserved performance on the Stroop on the other. This observation corroborates our finding that differences in PFC patients' performance emerge according to the type of inhibitory task administered. Importantly, our analyses indicated that differences in PFC performance are not mediated entirely by fluid intelligence. Instead, we found that they are associated with lesion location. Clearly lateralized damage to the right or left PFC is responsible for impairment on the two suppression measures of the Hayling or the Stroop respectively, despite the fact that both tasks require a verbal response. These findings are in broad agreement with previous literature suggesting a critical role of the right PFC for the Hayling suppression measures (e.g. Cipolotti et al., 2015a; Robinson et al., 2015) and of the left PFC for Stroop (e.g. Glascher et al., 2012; Tsuchida & Fellows, 2013; Geddes et al., 2014).

The lateralization findings of our Cognitive Investigations have been complemented and extended by the results of our VLSM Investigations. Our map of power distribution indicated a symmetrical left/right distribution for both tasks. We also found no significant difference in the lesion size between males and females or right and left frontal patients. Thus, it is unlikely that our VLSM results of lateralization for each task can be accounted for by differences in: unbalanced voxel power distribution, or interactions of lesions size with gender or laterality. VLSM analysis indicated that performance on both Suppression errors and Suppression RT2 was correlated with the integrity of a similar and relatively circumscribed region within the right lateral PFC, centred in the right lateral superior frontal gyrus and the right middle frontal gyrus. We previously documented, using a visual analysis of MRI/CT scans, that the right lateral frontal region is crucial for good performance on the Hayling. Thus, Cipolotti et al. (2015a) reported a marked impairment in suppression errors in a small sample of frontal patients with right lateral lesion. Robinson et al. (2015) found that patients with right lateral lesions (including lesions to middle and inferior gyrus only) were significantly impaired compared to HC on suppression errors. Patients with RIFG lesions were more likely to obtain impaired suppression scores than patients with LIFG. In our current study, controlling for the first time for fluid intelligence in performance on the two inhibitory tasks and adopting a different neuroimaging technique which makes no a priori

assumption about structure-function relationship, we confirmed the crucial role of the right lateral PFC on the two suppression measures of the Hayling. At this stage we would be cautious over conclusions concerning precise lesion localization within the right lateral PFC; of course, lesions to adjacent regions are inevitably correlated in any particular patient sample, restricting the exact interpretation of findings from VLSM.

Our VLSM findings for Stroop were in stark contrast with those for the Hayling. We found that performance on the Stroop was correlated with integrity of left lateral superior and middle frontal gyri. We had a good coverage of the ACC, yet we found no association between poor performance on the Stroop and damage to this area. Our ROI analysis found no significant difference in the performance of patients with lesions involving and not involving the ACC on the Stroop. Taken together our results are in line with recently published VLSM studies documenting that left PFC lesions are critical for the Stroop. However, our left lateral superior and middle frontal gyri result differs somewhat from findings in a number of other studies. Glascher and colleagues (2012) suggested that poor performance on the Stroop is associated with lesions in left dorsolateral PFC and assigned a role of inhibitory control to this region. An fMRI study also reported that the size of the behavioral Stroop interference effect was significantly correlated with left dorsolateral activity (Floden et al., 2011). Other studies have also implicated a more ventral left lateral frontal area in the Stroop. Thus, Tsuchida and Fellows (2013) using VLSM documented a left ventrolateral PFC effect for the Stroop. Similarly, Geddes and colleagues (2014) reported an exaggerated interference effect in the Stroop in a small sample of patients with left ventrolateral PFC damage (n=3) and suggested that this area was critical for interference resolution. Derrfuss and colleagues (2005) reported a common cluster of activation in the left inferior frontal junction for the Stroop. Again, given the limitations of VLSM, we would be cautious about exact localization of critical sites within the left PFC. Overall however, our cognitive and imaging findings for the Stroop support the view that the left PFC carries out a set of specific, localized functions relating to inhibition/interference control in this task. One possibility is that this arises from a failure to modulate speed-accuracy trade-offs appropriately, an important process in carrying out the Stroop (Kerns et al., 2004), where moving rapidly to a more accurate strategy activates left dorsolateral PFC (Vallesi et al., 2012; Vanderhasselt et al., 2009).

Our results allow us to conclude that lesion location, right or left PFC, is a critical factor in producing impairments not related to fluid intelligence on two inhibitory tasks

loading similarly on verbal control. This suggests that the two suppression measures of the Hayling and the Stroop are likely to assess dissociable components of executive functions, related to anatomically defined and lateralized PFC circuits. Moreover, our findings raise the possibility that the investigated core component of executive function - inhibition - may actually comprise qualitatively different forms with different neural substrates. In line with this, Kok (1999) reviewed a large body of behavioural and psychophysiological studies and suggested that there are multiple forms of inhibition, expressed in different ways and with distinct and interacting neuronal systems. Of course, it remains possible that the differences in performance we found between Hayling and Stroop and the effects of lesions may not be due to different inhibitory processes per se. Both our tasks require suppressing a dominant response, but also a variety of other complex processes. For example, for the Hayling, strategy implementation is needed to generate semantically unrelated alternatives sentences. In contrast, for the Stroop, the response set is very constrained and, as such, may require less strategy. For the Stroop also, processes involved in response conflict may play a greater role compared with the Hayling. The automatic response that needs to be suppressed is presented in the test stimuli (i.e. the incongruent colour-word). This may induce more conflict than in the case of the Hayling, where the response that needs to be suppressed is not visually or aurally presented. Future work is needed to better define the putative executive processes involved in these two tasks widely assumed to require inhibition.

One of the most influential views regarding the neuroanatomical underpinnings of inhibition suggests a critical role for the RIFG. Aron and colleagues (2014) and Robbins (2007) have proposed that right PFC damage leads to impairment in reactively suppressing inappropriate responses. Association between the RIFG and inhibitory processes has been supported by 'virtual lesion' evidence using transcranial magnetic stimulation (e.g., Chambers et al., 2006; Verbruggen et al., 2010) and by imaging and electrophysiological studies (for a review see Aron, 2011; but see Hampshire et al., 2010; Battaglia-Mayer et al., 2014, for an alternative perspective). Interestingly, Picton and colleagues (2007) provide contrasting evidence, suggesting that the left superior portion of BA6 is involved in the inhibition of responses in a go-no-go task. We had relatively good coverage of RIFG, yet our VLSM analyses found no association between damage to this area and poor performance on the Stroop. Moreover, our ROI analysis failed to find a significant difference in Stroop performance between patients with lesions involving the RIFG and patients with lesions not involving the RIFG. Notably, none of

the patients with lesions involving the RIFG had a frank impairment on this test. Thus, a clear discrepancy emerges when contrasting the RIFG findings for inhibition with our findings. Of course, the stop-signal and go-no-go tasks differ from the Stroop in several important dimensions such as the response modality. The Stroop, but not the other two tasks, also requires giving an alternate response rather than just activating an outright stopping mechanism. It is plausible that these important differences are implemented by different executive processes. Nonetheless, all these tasks require suppressing responses, an inhibitory role thought to be implemented by the RIFG (e.g. Aron et al., 2003). In our view, the unimpaired performance of our patients with RIFG lesions on the Stroop coupled with impaired performance following left PFC lesions is difficult to reconcile with a proposal that the RIFG is critical for inhibitory control in general (see for similar conclusions, Tsuchida & Fellows, 2013).

Our right PFC findings for the two suppression measures of the Hayling may be considered broadly consistent with the notion that damage to this region is associated with inhibition impairment (e.g. Hornberger & Bertoux, 2015). We have previously argued that the failure to attain an effective strategy may be one cause of suppression impairment on the Hayling. Burgess and Shallice (1996) argued that a failure in strategy-attainment was central to suppression impairment in the Hayling. If patients cannot generate a correct strategy, it is more difficult for them to inhibit pre-potent incorrect responses, resulting in a higher number of Suppression errors and elevated Suppression RT2. In keeping with this, we found that our right frontal patients made significantly more Suppression errors and had significantly elevated suppression RT2 compared with HC. Our VLSM analysis indicated an association between Suppression errors and right PFC damage. We previously reported that a high number of Suppression errors and a low number of correct responses compatible with a strategy are common in right lateral but not orbitofrontal patients (Cipolotti et al., 2015a). We suggest that, in the Hayling, impairments after right PFC lesions reflect not a domain-general inhibitory deficit, but a specific difficulty in suppression of a dominant response by virtue of generating an alternative, task-appropriate strategy.

The existence of dissociations such as those reported in our study are difficult to reconcile with any simple account based on left hemisphere dominance for verbal tasks. We documented contrasting right and left PFC involvement for two verbal tasks requiring verbal responses. Similarly, our findings are difficult to explain in terms of a general failure to shape performance by task goals. According to this view a task goal

impairment would lead to deficits in many tasks, such as the Hayling, go no-go, or stop-signal tasks (Hornberger & Bertoux, 2015). For example, a failure to maintain the task goal in the Hayling would lead to inhibition/suppression deficits as a result of erratic responding. In support of this position, a recent meta-analysis of functional neuroimaging executive tasks showed that the RIFG and the right anterior insula are activated in several executive tasks requiring task goal maintenance (e.g. Cieslik et al., 2015). However, general task goal maintenance is also required for the Stroop (e.g. Cohen & Servan-Schreiber, 1992; West & Baylis, 1998), yet performance on this task, compared to Hayling, was dependent more on left than right PFC. It seems that a general failure in control of behaviour by task goals cannot be easily reconciled with the differentially lateralized functional organization we documented.

As far as we are aware, our study represents the first investigation of the complex relationship between performances on two inhibitory tasks requiring verbal responses, and fluid intelligence in patients with focal unilateral frontal lesions. Our findings are clinically relevant and contribute to the debate regarding the PFC contribution to inhibition and the functional organization of the PFC. Of course, there are a number of important methodological limitations to consider. For example, VLSM has been considered a tremendous step forward when attempting to map brain circuits to specific functions (e.g. Karnath & Smith, 2014). However, it cannot overcome intrinsic data limitations due to correlated damage to adjacent regions, coupled with differential power arising from differences in lesion distribution (Mah et al., 2014). Varying chronicity may also contribute to varying outcome across studies; in our study, mean chronicity of lesions was only 8.87 months, compared e.g. to 3.2 and 4.5 years in the studies of Aron et al. (2004) and Tsuchida and Fellows (2013). Some lesion studies have reported that men and women can show differing lateralization in their neuroanatomy, particularly in the ventral medial PFC (Piefke et al., 2005; Stevens & Hamann, 2012; Sutterer et al., 2015; Tranel et al., 2005). Due to our patient sample size, we could not formally assess whether gender may have had an effect on our findings. However, we found no significant gender differences in the behavioural performance of the right frontal patients on the Suppression Errors (p=.570) nor on the RT2 (p=.912), the left frontal patients on the Stroop (p=.889), or in the lesion size between right and left frontal patients used in the VLSM analysis.

In conclusion, we would like to propose that there are several processes controlled by anatomically separable systems involved in response inhibition tasks (see

Kok, 1999, for a similar view). The left PFC cortex may be involved in inhibitory control of pre-potent responses and in response conflict or in modulations of speed-accuracy trade-off. The right PFC may be involved in inhibiting dominant strategies and in implementing new strategies. Our findings emphasize the importance of continuing lesion studies to further the understanding of the PFC contribution to executive processes such as inhibition. Moreover, our findings have clinical implications for the diagnosis and treatment of disinhibition impairments. They highlight the need to assess inhibition in frontal patients using a variety of tests and to develop different types of treatments, since we demonstrated that two widely used tests, traditionally thought to require inhibition, rely on differently lateralized left or right PFC circuits.

Funding

This work was supported by the Welcome Trust Grant (089231/A/09/Z). This work was undertaken at UCLH/UCL, which received a proportion of funding from the Department of Health's National Institute for Health Research Biomedical Research Centre's funding scheme.

Legend to Figures

Figure 1. Lesion maps of PFC patients, overlaid onto a T1-weighted brain image in Montreal Neurological Institute space, in 3D views and in axial slices. The colour bar indicates the number of patients with lesion in a given brain area. Warmer areas indicate areas of greater lesion overlap.
Abbreviations: L=left; R= right.

Figure 2. Map of power distributions for the Hayling (panel A) and Stroop (panel B). Power values range from 0.1 (dark blue) to 0.4 (light blue/green). The power map reflects the probability to detect a true relationship between damage in a certain region and a behavioural score of interest. Abbreviations: L=left; R= right
Abbreviations: L=left; R= right.

Figure 3. Mean z-scores of left and right frontal patients on Suppression Errors, Suppression RT2 and Stroop Colour-Word Test, after accounting for RAPM.
Abbreviations: RAPM=Ravens Advanced Progressive Matrices

Figure 4. Results of voxel lesion-symptom mapping analyses (VLSM), overlaid onto a T1-weighted brain image in Montreal Neurological Institute space, in 3D views and in axial slices. Voxels in red show the area found to be significant (p<0.05 FWE-corrected at cluster level) for (A) Hayling Suppression error scores, (B) Hayling Suppression RT2 scores, (C) Stroop scores.
Abbreviations: L=left; R= right

References

Aron, A.R., 2011. From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. Biol. Psychiatry 69, e55–68.

Aron, A.R., Fletcher, P.C., Bullmore, E.T., Sahakian, B.J., Robbins, T.W., 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nat. Neurosci. 6, 115–116.

Aron, A.R., Monsell, S., Sahakian, B.J., Robbins, T.W., 2004a. A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. Brain 127, 1561–1573.

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2014. Inhibition and the right inferior frontal cortex: one decade on. Trends Cogn. Sci. 18, 177–185.

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004b. Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. 8, 170–177.

Arthur, W., Day, D.V., 1994. Development of a Short form for the Raven Advanced Progressive Matrices Test. Educ. Psychol. Meas. 54, 394–403.

Badre, D., 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn. Sci. 12, 193–200.

Badre, D., D'Esposito, M., 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? Nat. Rev. Neurosci. 10, 659–69.

Baird, A., Dewar, B.-K., Critchley, H., Gilbert, S.J., Dolan, R.J., Cipolotti, L., 2006. Cognitive functioning after medial frontal lobe damage including the anterior cingulate cortex: a preliminary investigation. Brain Cogn. 60, 166–175.

Bates, E., Wilson, S.M., Saygin, A.P., Dick, F., Sereno, M.I., Knight, R.T., Dronkers, N.F., 2003. Voxel-based lesion-symptom mapping. Nat. Neurosci. 6, 448–450.

Battaglia-Mayer, A., Buiatti, T., Caminiti, R., Ferraina, S., Lacquaniti, F., Shallice, T., 2014. Correction and suppression of reaching movements in the cerebral cortex: Physiological and neuropsychological aspects. Neurosci. Biobehav. Rev. 42, 232–251.

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

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. Psychol. Rev. 108, 624–652.

Burgess, P.W., Shallice, T., 1996. Response suppression, Initiation and strategy use following frontal lobe lesions. Neuropsychologia 34, 263–272.

Carter, C.S., Macdonald, A.M., Botvinick, M., Ross, L.L., Stenger, V.A., Noll, D., Cohen, J.D., 2000. Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 1944–1948.

Chambers, C.D., Bellgrove, M.A., Stokes, M.G., Henderson, T.R., Garavan, H., Robertson, I.H., Morris, A.P., Mattingley, J.B., 2006. Executive "brake failure" following deactivation of human frontal lobe. J. Cogn. Neurosci. 18, 444–455.

Cieslik, E.C., Mueller, V.I., Eickhoff, C.R., Langner, R., Eickhoff, S.B., 2015. Three key regions for supervisory attentional control: evidence from neuroimaging meta-analyses. Neurosci. Biobehav. Rev. 48, 22–34.

Cipolotti, L., Healy, C., Chan, E., Bolsover, F., Lecce, F., White, M., Spano, B., Shallice, T., Bozzali, M., 2015a. The impact of different aetiologies on the cognitive performance of frontal patients. Neuropsychologia 68, 21–30.

Cipolotti, L., Healy, C., Spano, B., Lecce, F., Biondo, F., Robinson, G., Chan, E., Duncan, J., Shallice, T., Bozzali, M., 2015b. Strategy and suppression impairments after right lateral prefrontal and orbito-frontal lesions. Brain e1, 1–4.

Cohen, J.D., Servan-Schreiber, D., 1992. Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. Psychol. Rev. 99, 45–77.

De Pisapia, N., Braver, T.S., 2006. A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. Neurocomputing 69, 1322–1326.

Demakis, G.J., 2004. Frontal Lobe Damage and Tests of Executive Processing: A Meta-Analysis of the Category Test, Stroop Test, and Trail-Making Test. J. Clin. Exp. Neuropsychol. 26, 441–450.

Derrfuss, J., Brass, M., Neumann, J., Von Cramon, D.Y., 2005. Involvement of the inferior frontal junction in cognitive control: Meta-analyses of switching and stroop studies. Hum. Brain Mapp. 25, 22–34.

Drewe, E., 1975. Go-no go learning after frontal lobe lesions in humans. Cortex A J. Devoted to Study Nerv. Syst. Behav. 11, 8–16.

Duncan, J., 2001. An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. 2, 820–829.

Duncan, J., Burgess, P., Emslie, H., 1995. Fluid intelligence after frontal lobe lesions. Neuropsychologia 33, 261–268.

Eriksen, B. a., Eriksen, C.W., 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept. Psychophys. 16, 143–149.

Fellows, L.K., Farah, M.J., 2005. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. Cereb. Cortex 15, 58–63.

Floden, D., Vallesi, A., Stuss, D.T., 2011. Task context and frontal lobe activation in the Stroop task. J. Cogn. Neurosci. 23, 867–879.

Friedman, N.P., Miyake, A., 2004. The relations among inhibition and interference control functions: a latent-variable analysis. J. Exp. Psychol. Gen. 133, 101–135.

Geddes, M.R., Tsuchida, A., Ashley, V., Swick, D., Fellows, L.K., 2014. Material-specific interference control is dissociable and lateralized in human prefrontal cortex. Neuropsychologia 1–10.

Glascher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., Paul, L.K., Tranel, D., 2012. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proc. Natl. Acad. Sci. 109, 14681–14686.

Glascher, J., Tranel, D., Paul, L.K., Rudrauf, D., Rorden, C., Hornaday, A., Grabowski, T., Damasio, H., Adolphs, R., 2009. Lesion Mapping of Cognitive Abilities Linked to Intelligence. Neuron 61, 681–691.

Hampshire, A., Chamberlain, S.R., Monti, M.M., Duncan, J., Owen, A.M., 2010. The role of the right inferior frontal gyrus: inhibition and attentional control. Neuroimage 50, 1313–1319.

Hornberger, M., Bertoux, M., 2015. Right lateral prefrontal cortex—specificity for

inhibition or strategy use? Brain 138, 833–5.

Hornberger, M., Geng, J., Hodges, J.R., 2011. Convergent grey and white matter evidence of orbitofrontal cortex changes related to disinhibition in behavioural variant frontotemporal dementia. Brain 134, 2502–2512.

Karnath, H.-O., Smith, D.V., 2014. The next step in modern brain lesion analysis: multivariate pattern analysis. Brain 137, 2405–2407.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior cingulate conflict monitoring and adjustments in control. Science 303, 1023–1026.

Kimberg, D.Y., Coslett, H.B., Schwartz, M.F., 2007. Power in voxel-based lesion-symptom mapping. J. Cogn. Neurosci. 19, 1067–1080.

Kimberg, D.Y., Farah, M.J., 1993. A unified account of cognitive impairments following frontal lobe damage: the role of working memory in complex, organized behavior. J. Exp. Psychol. Gen.

Knutson, K.M., Rakowsky, S.T., Solomon, J., Krueger, F., Raymont, V., Tierney, M.C., Wassermann, E.M., Grafman, J., 2013. Injured brain regions associated with anxiety in Vietnam veterans. Neuropsychologia 51, 686–694.

Kok, A., 1999. Varieties of inhibition: manifestations in cognition, event-related potentials and aging. Acta Psychol. (Amst). 101, 129–158.

Logan, G.D., 1994. On the ability to inhibit thought and action: A users' guide to the stop signal paradigm, in: Inhibitory Processes in Attention, Memory, and Language. pp. 189–239.

Mah, Y.-H., Husain, M., Rees, G., Nachev, P., 2014. Human brain lesion-deficit inference remapped. Brain 137, 2522–31.

McKenna, P., Warrington, E., 1980. The Graded Naming Test. NFER-Nelson, Windsor.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 24, 167–202.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, a H., Howerter, A., Wager, T.D., 2000. The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. Cogn. Psychol. 41, 49–100.

Nelson, H.E., 1982. National Adult Reading Test. NFER-Nelson, Windsor.

Obonsawin, M.C., Crawford, J.R., Page, J., Chalmers, P., Cochrane, R., Low, G., 2002. Performance on tests of frontal lobe function reflect general intellectual ability.

Neuropsychologia 40, 970–977.

Parvizi, J., 2012. Disinhibition: More than a misnomer. Soc. Neurosci. 7, 311–316. Perret, E., 1974. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia 12, 323–330.

Petrides, M., 2005. Lateral prefrontal cortex: architectonic and functional organization. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 781–795.

Picton, T.W., Stuss, D.T., Alexander, M.P., Shallice, T., Binns, M. a., Gillingham, S., 2007. Effects of focal frontal lesions on response inhibition. Cereb. Cortex 17, 826–838.

Piefke, M., Weiss, P.H., Markowitsch, H.J., Fink, G.R., 2005. Gender differences in the functional neuroanatomy of emotional episodic autobiographical memory. Hum. Brain Mapp. 24, 313–324.

Raven, J.C., 1976. Manual for the Advanced Progressive Matrices: Set 1.

Reitan, R.M., 1992. Trail Making Test: Manual for administration and scoring. Reitan Neuropsychology Laboratory, South Tucson, AZ.

Robbins, T.W., 2007. Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 362, 917–932.

Robinson, G.A., Cipolotti, L., Walker, D.G., Biggs, V., Bozzali, M., Shallice, T., 2015. Verbal suppression and strategy use: a role for the right lateral prefrontal cortex? Brain 138, 1084–1096.

Robinson, G.A., Shallice, T., Bozzali, M., Cipolotti, L., 2012. The differing roles of the frontal cortex in fluency tests. Brain 135, 2202–2214.

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., Duncan, J., 2010. Executive function and fluid intelligence after frontal lobe lesions. Brain 133, 234–247.

Shallice, T., Gillingham, S.M., 2012. On Neuropsychological Studies of Prefrontal Cortex: The Robbia Approach, in: Stuss, D.T., Knight, R.T. (Eds.), Principles of Frontal Lobe Function. Oxford University Press, Oxford.

Shallice, T., Stuss, D.T., Picton, T.W., Alexander, M.P., Gillingham, S., 2008. Mapping task switching in frontal cortex through neuropsychological group studies. Front. Neurosci.

Stevens, J.S., Hamann, S., 2012. Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. Neuropsychologia 50, 1578–1593.

Stroop, J.R., 1935. Stroop color word test, J. Exp. Physiol.

Stuss, D.T., Alexander, M.P., 2007. Is there a dysexecutive syndrome? Philos.trans.r.soc.l. B Biol.Sci. 362, 901–915.

Stuss, D.T., Floden, D., Alexander, M.P., Levine, B., Katz, D., 2001. Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. Neuropsychologia 39, 771–786.

Stuss, D.T., Levine, B., 2002. Adult Clinical Neuropsychology: Lessons from Studies of the Frontal Lobes. Annu. Rev. Psychol. 53, 401–433.

Sutterer, M.J., Koscik, T.R., Tranel, D., 2015. Sex-related functional asymmetry of the ventromedial prefrontal cortex in regard to decision-making under risk and ambiguity. Neuropsychologia 75, 265–273.

Swick, D., Turken, A.U., 2002. Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. Proc. Natl. Acad. Sci. U. S. A. 99, 16354–16359.

Torso, M., Serra, L., Giulietti, G., Spanò, B., Tuzzi, E., Koch, G., Caltagirone, C., Cercignani, M., Bozzali, M., 2015. Strategic lesions in the anterior thalamic radiation and apathy in early Alzheimer's disease. PLoS One 10, e0124998.

Tranel, D., Damasio, H., Denburg, N.L., Bechara, A., 2005. Does gender play a role in functional asymmetry of ventromedial prefrontal cortex? Brain 128, 2872–2881.

Trenerry, M., Crosson, B., DeBoe, J., & Leber, W., 1989. Stroop Neuropsychological Screening Test manual. Psychological Assessment Resources (PAR), Odessa, FL.

Tsuchida, A., Fellows, L.K., 2013. Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. Cortex 49, 1790–1800.

Turner, M.S., Cipolotti, L., Yousry, T., Shallice, T., 2007. Qualitatively different memory impairments across frontal lobe subgroups. Neuropsychologia 45, 1540– 1552.

Vallesi, A., Mcintosh, A.R., Crescentini, C., Stuss, D.T., 2012. fMRI investigation of speed-accuracy strategy switching. Hum. Brain Mapp. 33, 1677–1688.

Vanderhasselt, M.-A., De Raedt, R., Baeken, C., 2009. Dorsolateral prefrontal cortex and Stroop performance: tackling the lateralization. Psychon. Bull. Rev. 16, 609–612.

Vendrell, P., Junqué, C., Pujol, J., Jurado, M. a, Molet, J., Grafman, J., 1995. The role of prefrontal regions in the Stroop task. Neuropsychologia 33, 341–352.

Verbruggen, F., Aron, A.R., Stevens, M.A., Chambers, C.D., 2010. Theta burst 32

stimulation dissociates attention and action updating in human inferior frontal cortex. Proc. Natl. Acad. Sci. 107, 13966–13971.

Volle, E., De Lacy Costello, A., Coates, L.M., McGuire, C., Towgood, K., Gilbert, S., Kinkingnehun, S., McNeil, J.E., Greenwood, R., Papps, B., Van Den Broeck, M., Burgess, P.W., 2012. Dissociation between verbal response initiation and suppression after prefrontal lesions. Cereb. Cortex 22, 2428–2440.

West, R., Baylis, G.C., 1998. Effects of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. Psychol. Aging 13, 206–217.

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., Roca, M., Antoun, N., Manes, F., Duncan, J., 2010. Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. Proc. Natl. Acad. Sci. U. S. A. 107, 14899–14902.

Zook, N.A., Davalos, D.B., DeLosh, E.L., Davis, H.P., 2004. Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. Brain Cogn. 56, 286–292.

Table 1. Demographic and Cognitive Data: Frontal Patients and Healthy Comparison

	Frontal Patients	**Healthy Comparison**
	N = 30	**N = 60**
Age (years)	47.40	46.41
(SD)	(15.52)	(14.85)
Education (years)	13.40	13.53
(SD)	(2.71)	(2.72)
NART IQ	108.03	111.25
(SD)	(9.50)	(7.29)
Gender (Male/Female)	20/10	35/25
Time Between Damage and Assessment (months)	8.86	N/A
(SD)	(20.84)	N/A
RAPM (Correct Responses/12)	**7.70***	9.26
(SD)	(2.38)	(1.85)
GNT (Items Correctly Identified/30)	20.56	21.93
(SD)	(4.93)	(3.37)
Fluency S (Words Produced)	14.46	16.20
(SD)	(6.10)	(5.00)
TMT-A (Seconds)	**37.45***	30.95
(SD)	(14.66)	(10.65)

Legend. Scores with significant *p*-values are in bold; *$p < 0.05$ compared with Healthy Comparison;

(SD): Standard Deviation;

N/A: not applicable;

NART: National Adult Reading Test;

RAPM: Ravens Advanced Progressive Matrices;

GNT: Graded Naming Test;

TMT-A: Trail Making Test – Part A

Table 2. Mean Scores on Part 2 of the Hayling and Stroop Colour-Word Tests: Frontal patients and Healthy Comparison

	Frontal Patients	Healthy Comparison
Suppression Error SS	**4.80***	6.05
(SD)	(2.60)	(1.80)
Suppression RT2 SS	**4.73***	6.00
(SD)	(1.76)	(0.63)
Stroop (No. of colours named in 2 minutes)	**94.59***	114.49
(SD)	(31.38)	(20.30)

Legend. Scores with significant *p*-values are in bold; *$p < 0.05$ compared with Healthy Comparison;

SS: Scaled Score;

(SD): Standard Deviation

RT: reaction time

Table 3. Correlations - Part 2 of the Hayling and Stroop Colour-Word Tests

	Correlations		Correlations adjusted for RAPM	
	r	***p*-value**	**r**	***p*-value**
Suppression Errors **Suppression RT2**	.605	**< .001**	.435	**.038**
Suppression Errors **Stroop Colour-Word Test**	.527	**.005**	.311	.170
Suppression RT2 **Stroop Colour-Word Test**	.524	**.005**	.163	.479

Legend. *: $p < .05$;

RAPM: Ravens Advanced Progressive Matrices

RT: reaction time

Figure 1

Click here to download high resolution image

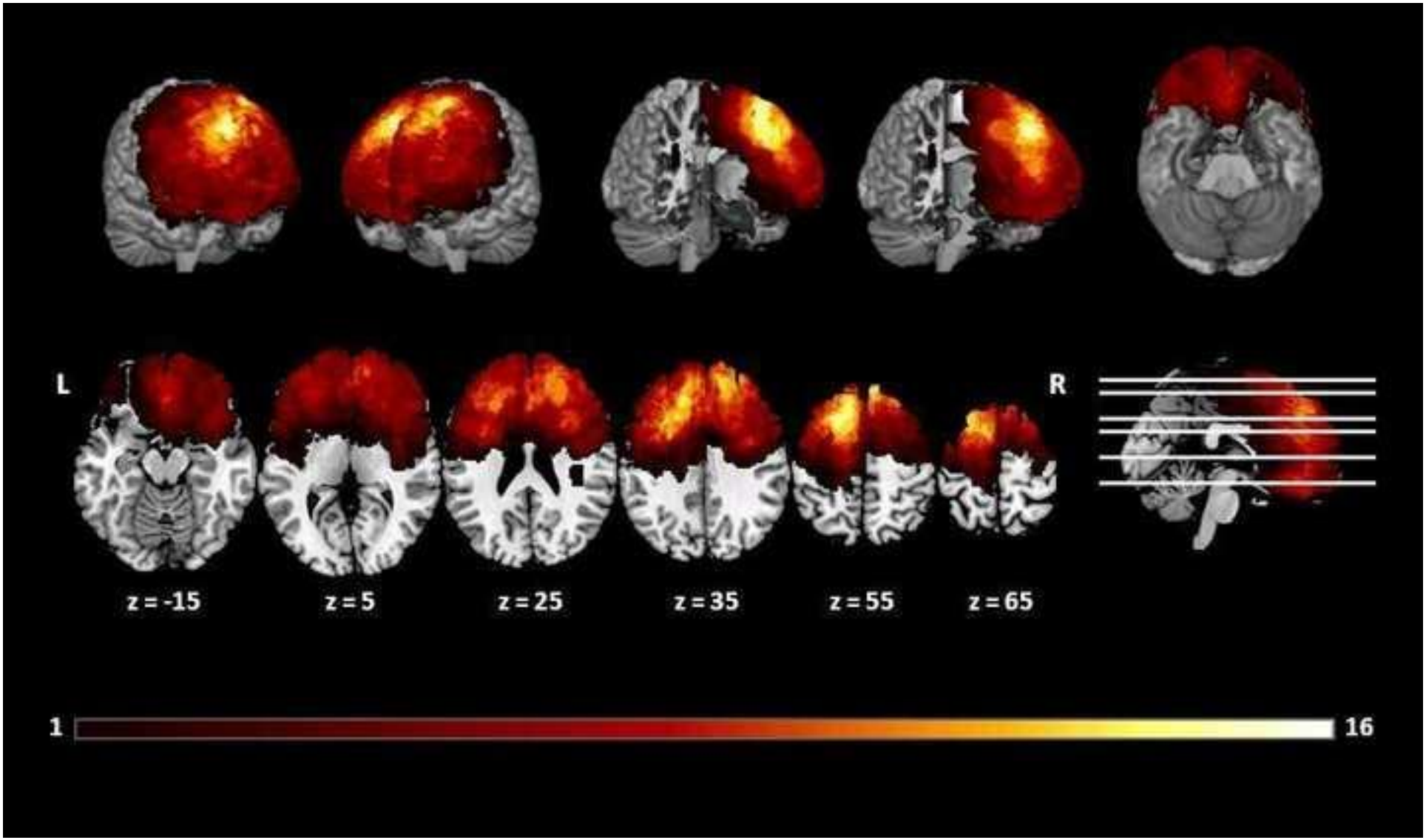

Figure 2

Click here to download high resolution image

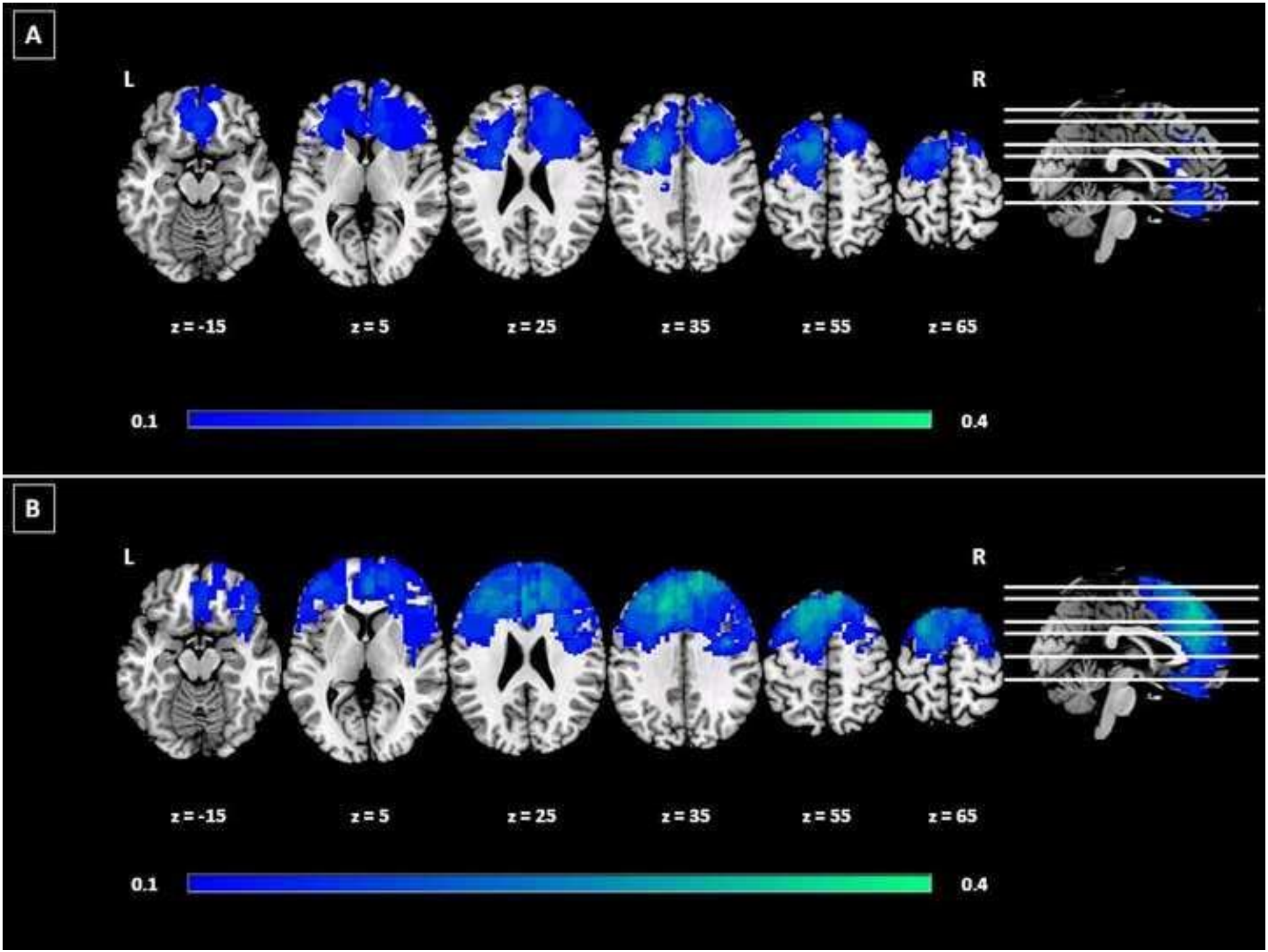

Figure 3

Click here to download high resolution image

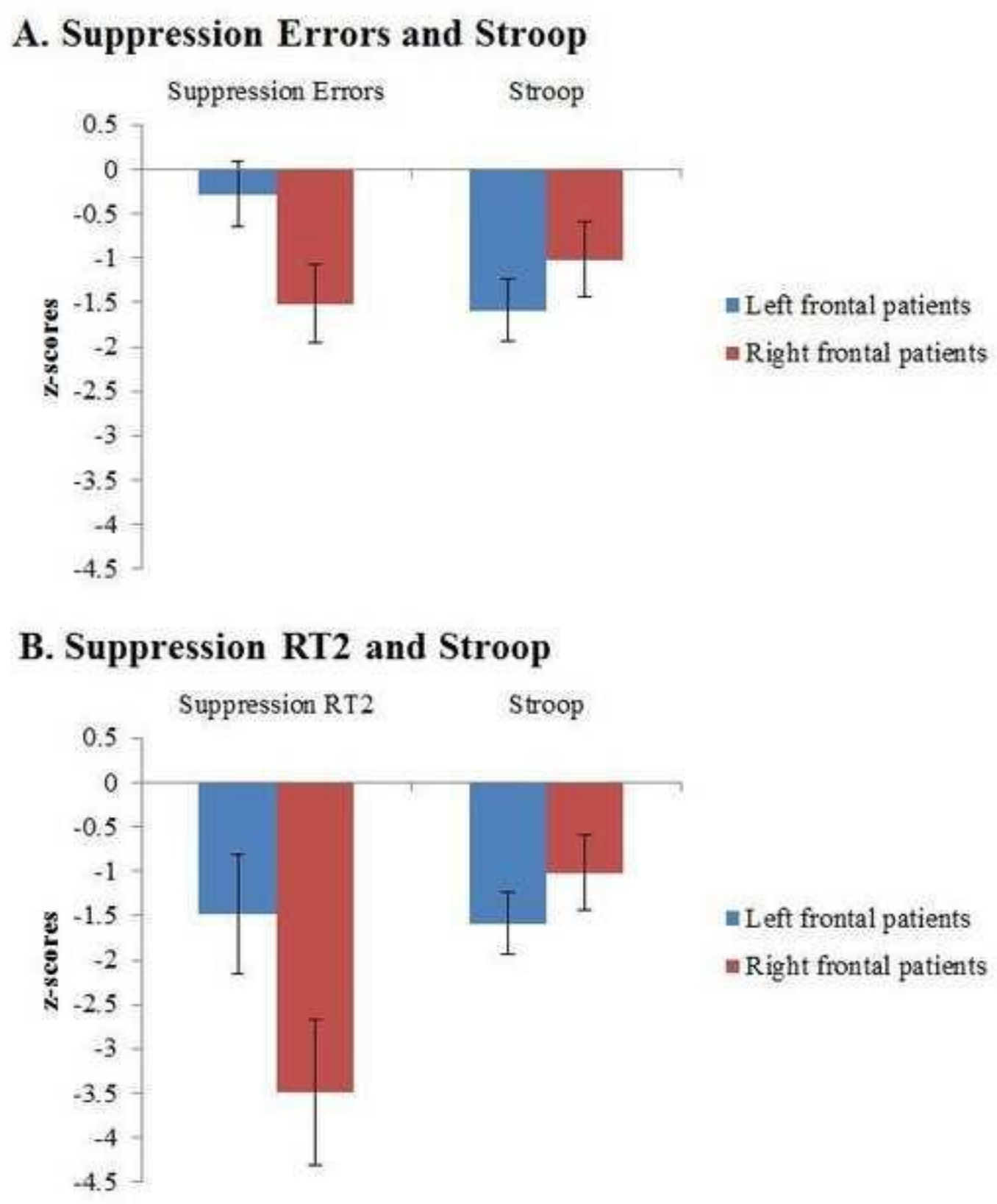

Figure 4

Click here to download high resolution image