Restorative Neurology and Neuroscience 27 (2009) 493–506
DOI 10.3233/RNN-2009-0511
IOS Press

Mesulam's frontal lobe mystery re-examined

Paul W. Burgess[a,*], Nick Alderman[b], Emmanuelle Volle[a], Roland G. Benoit[a] and Sam J. Gilbert[a]
[a]*Institute of Cognitive Neuroscience, UCL* (*University College London*)*, 17 Queen Square, London WC1N 3AR, UK*
[b]*Kemsley National Centre for Neurorehabilitation, St Andrew's Hospital, Northampton NN1 5DG, UK*

Abstract. *Purpose*: Mesulam's (1986) mystery is that some patients with frontal lobe damage may show no cognitive impairment according to traditional office-based assessment procedures, yet nevertheless show marked cognitive handicap in everyday life. Mesulam suggested that "the office setting may introduce sufficient external structure to suppress some of these behavioral tendencies" (p. 322). We ask if it is indeed the office setting that is the problem, or whether it is that traditional assessments do not measure the full range of cognitive functions supported by prefrontal cortex.
Method: Neuropsychological case series study and review.
Results: Traditional methods for assessing cognitive deficits following frontal lobe damage typically do not measure the full range of deficits that can occur. In particular, rostral prefrontal cortex supports functions which are not routinely assessed yet are crucial to competent everyday life performance. These include meta-memory functions (e.g. context and source memory), complex behavioural co-ordination (e.g. prospective memory and multitasking), and mentalizing.
Conclusions: New clinical assessment procedures are required urgently. These could be based, in principle, both on recent experimental findings from cognitive neuroscience, and observation of behaviour outside office settings. These procedures could then be administered in an office setting.

1. Introduction

> *"Some patients with sizable frontal lobe lesions may have routine neurological and neuropsychological examinations that are quite unremarkable. This creates a problem in the assessment of these patients, especially since the behavioural derangements – which sometimes constitute the only salient features – are also too complex to test in the office . . . the office setting may introduce sufficient external structure to suppress some of these behavioral tendencies...The clinical adage that judgement and complex comportment cannot be tested in the office is particularly pertinent to the evaluation of patients with frontal lobe damage." [Mesulam* (*1986*) *pp. 321–322.]*

Marcel Mesulam's influential editorial paper captured brilliantly the essence of a frequently discussed clinical problem in neurology. This is that some people who have suffered frontal lobe damage clearly have impairments in "everyday life" settings, but show little or no impairment on the wide range of formal clinical tests of cognition available to the neuropsychologist or neurologist. This paper examines what we now know about this phenomenon, more than 20 years after Mesulam's observation, and asks if the views he expressed about the causes of the problem might now need modification.

Mesulam was not suggesting that all patients with frontal lobe damage, or even a large proportion of them, would show this apparently quite specific pattern. Indeed, he provides a list of the remarkably broad range of behavioural alterations or "dysexecutive symptoms" that have been associated with frontal lobe involvement: puerility, profanity, slovenliness, facetiousness, irresponsibility, grandiosity, irascibility; loss of spontaneity, curiosity and initiative, with apathetic blunting of feeling, drive, mentation and behaviour; "erosion of" foresight, judgement, insight; disturbances of the ability to delay gratification and capacity for remorse, and in abstract reasoning, creativity, problem-solving, and

*Corresponding author: Prof. Paul. W. Burgess, Institute of Cognitive Neuroscience, UCL (University College London), 17 Queen Square, London WC1N 3AR, UK. Tel.: +44 207 679 1177; Fax: +44 207 679 1177; E-mail: p.burgess@psychol.ucl.ac.uk.

mental flexibility; problems with jumping to premature conclusions and concrete or stimulus-bound behaviour; difficulties with the orderly planning and sequencing of complex behaviours, the ability to attend to several components simultaneously and alter "flexibly" the focus of one's attention; difficulties in grasping the context and gist of a complex situation, with resistance to distraction, the ability to follow multistep instructions, and deficits in inhibition of inappropriate responses and the ability to sustain behaviour without perseveration.

Looking through this list however, one cannot fail to be struck by the degree to which the behaviours or deficits described are *not* assessed by traditional clinical tests of the forms of cognition thought to be subserved by the frontal lobes ("executive functions"). There are, to our knowledge, no formal tests for "puerility, profanity, slovenliness facetiousness, irresponsibility, grandiosity, irascibility" nor a majority of the other characteristics mentioned.

This is perhaps not surprising. As Burgess et al. (2006) point out, traditional tests of executive function (e.g. Wisconsin Card Sorting Test (WCST), verbal fluency, Stroop and others) were not originally invented for the purpose of investigating executive functions in patient populations, and in many cases there has been little or no substantial modification to these procedures over the intervening years to accommodate this new application. Moreover, they were developed at a time when the theoretical understanding of executive function was in its infancy, and the conceptual frameworks in existence were very different from today. Indeed, even after many years of investigation we still know little about what these tests actually measure, or the range of activities that require the same processes. This is not helped by the fact that the tests are typically not good "models of the world" (i.e. the situation they present to the participant is not like those usually encountered in everyday life). So there will always be a question as to how relevant the test results can be in reflecting behaviour "outside the office".

Perhaps one might suggest that a form of scientific evolutionary selection might have occurred over the last 100 years, which has led to the circumstance where the most commonly used clinical assessment procedures are also those that tap some mental processing related to the most frequent symptoms that Mesulam describes. In this case there might be a rough correspondence over large groups of people between those who fail the tests and those who show the most obvious symptoms outside the office. However it need not follow that there is transparency "within category". In other words it need not be that someone who fails a test of "abstract reasoning" will more likely show behaviours described as "abstract reasoning" failures in other circumstances. Our extreme lack of understanding in this field means that the characteristics of these situations presented to the participant are currently theoretically and experimentally underspecified. We do not have defining features for a test of "abstract reasoning" that will reliably distinguish it from, say, a test of "planning". This is also true of most if not all of the constructs that Mesulam evokes above.

Adding to these difficulties, Mesulam does not specify in detail what are the particular behavioural characteristics of those patients whose difficulties "outside the office" are not well reflected by their performance on tests within it. Thus it is not clear whether the assessment problem he describes is caused (a) by the office situation and the constraints that imposes, or (b) simply because the correct tests are not given, or do not exist (or at least did not 20 years ago).

It is this second possibility that is examined here. We will suggest that there are many mental abilities supported, at least in part, by the frontal lobes that were not known at the time that Mesulam was writing. (Or, at least, that the experimental evidence for them was at that time too scant for its inclusion in a short Editorial.) Moreover, we propose that to a large degree this was because the functions of the largest single cytoarchitectonic subregion of human prefrontal cortex, rostral prefrontal cortex (incorporating Brodmann's Area 10; also referred to by other groups as "anterior PFC", "frontopolar cortex", or "the frontal pole") have, up until the last ten years or so, not been investigated. However we are perhaps now in a position to go some way towards solving Mesulam's mystery.

2. Everyday life problems despite good neuropsychological test performance: A brief history.

The kind of pattern that Mesulam highlighted (behavioural impairments in everyday life disproportionate to failures on formal examination) had received close investigation as early as fifty years before his article. Penfield and Evans (1935) described the symptoms that Penfield's sister was experiencing after the removal of a right frontal glioma: "She had planned to get a simple supper for one guest and four members of her family. She looked forward to it with pleasure and had the whole day for preparation. When the appointed hour arrived she was in the kitchen, the food was all there, one or two things were on the stove, but the salad was not ready, the meat had not been started and she was distressed and confused by her long continued effort alone".

This impairment was not a reflection of serious disabilities in basic cognitive systems (e.g. classic dense amnesia, visuo-spatial/perceptual or agnosic problems, disorders of motor control and so forth). These cognitive functions were generally unimpaired, as was the case for similar cases that were soon reported (e.g. Brickner, 1936; Ackerly and Benton, 1947). These cases established, at least on the grounds of clinical observation alone, that this kind of behavioural disorganisation can be seen in the absence of impairments in basic cognitive systems.

However it was not until the mid-1980s that an attempt was made to characterise the nature of the critical cognitive deficit underpinning this disorder. Eslinger and Damasio (1985) described the case of EVR, who had undergone surgical removal of a large bilateral frontal meningioma. At the time of his operation EVR was a financial officer with a small company and a respected member of his community. He was married and the father of two children; his brothers and sisters considered him a role model and a natural leader. After the operation however, EVR lost his job, went bankrupt, was divorced by his wife, and moved in with his parents. He subsequently married a prostitute and was divorced again within two years. Extensive psychological evaluations found no deficit; in fact, he was superior or above average on most tests (e.g., Verbal IQ of 125; Performance IQ of 124; no difficulty on Wisconsin Card Sorting Test). He was also able to discuss intelligently matters such as the economy, foreign affairs, financial matters, or moral dilemmas. Despite these normal findings, EVR was often unable to make simple everyday decisions, such as which toothpaste to buy, what restaurant to go to, or what to wear. He would instead make endless comparisons and contrasts, often being completely unable to come to a decision at all. Further, Eslinger and Damasio report: "... it was as if he forgot to remember short- and intermediate-term goals ... "(1985, p. 1737). This may be characterised as a failure of "prospective memory", the ability to encode delayed intentions, and act on those intentions when the appropriate time arrives. Although this domain had received scant attention at the time of Mesulam's (1986) editorial, there has been a proliferation of studies investigating this important topic in recent years (see e.g. Brandimonte et al., 1996; Kliegel et al., 2008 for overviews of the field). We will return to the topic of prospective memory below.

Eslinger and Damasio's important paper was the first convincing empirical demonstration that this level of behavioural disorganisation could occur in the context of intact intellect, and intact performance on some tests traditionally thought to be sensitive to deficits in "frontal lobe" executive functions. This group went on to explore an explanation for some of the concomitant emotional changes displayed by patients like EVR in terms of an acquired insensitivity to future consequences (the "somatic marker hypothesis"; e.g. Bechara et al., 1994), and developed the Iowa Gambling Task as an experimental measure of this propensity. This has provoked considerable interest and debate about the issue of emotional changes following frontal lobe damage (especially to ventromedial PFC; see e.g. Bechara et al., 2005), and has opened up new possible explanations, and potential assessment procedures, for behavioural abnormalities that would not have been available to Mesulam.

Meanwhile, Burgess and colleagues had also been working with patients like EVR, who appeared to be intact on traditional neuropsychological testing but who showed behavioural organization difficulties in everyday life. They addressed the issue of formal quantification of the problems experienced by these people in everyday life, since up to that point the literature had only reported descriptions of them.

The first of these studies was published five years after Mesulam's article. Shallice and Burgess (1991) presented three cases who had all suffered frontal lobe damage following traumatic brain injury. All three had no significant impairment on formal tests of perception, language and intelligence and two performed well on a variety of traditional tests of executive function. Indeed, one of these cases (AP) was probably the best example of the syndrome so far reported (this case was later called "NM" by Metzler and Parkin (2000). AP had sustained an open head injury in a road-traffic accident when he was in his early twenties. The injury caused a virtually complete removal of the rostral prefrontal cortex bilaterally plus damage to surrounding regions. On standard neuropsychological measures of intellectual functioning, memory, perception and even traditional tests of executive function, AP performed within the superior range.

However, AP did show cognitive impairment in other situations (Shallice and Burgess, 1991; Metzler and Parkin, 2000). The most noticeable of these in everyday life was a marked multitasking and prospective memory problem. This manifested itself as tardiness and disorganisation, the severity of which ensured that despite his excellent intellect and social skills, he never managed to make a return to work at the level he had enjoyed pre-morbidly. Shallice and Burgess (1991) invented two new tests of multitasking to assess these problems. The first of these tests, called the "Multiple Errands Test" was a real-life multitasking test carried out in shopping precinct. Participants have to complete a number of tasks, principally involving shopping in an unfamiliar shopping precinct, whilst following a set of rules (e.g. no shop should be entered other than to buy something). The tasks vary in terms of complexity (e.g. buy a small brown loaf vs. discover the exchange rate of a particular foreign currency yesterday), and there are a number of "hidden" problems in the tasks that have to be appreciated and the possible course of action evaluated (e.g. one item asks that participants write and send a postcard, yet they are given no pen, and although they cannot use anything not bought on the street to help them, they are also told that need to spend as little money as possible). In this way, the task is quite "open-ended" or "ill-structured" (i.e. there are many possible courses of action, and it is up to the individual to determine for themselves which one they will choose).

The second task that Shallice & Burgess invented was a more controlled experimental task (the "Six Element Test"). This required subjects to swap efficiently between 3 simple subtasks, each divided into two sections within 15 minutes, whilst following some arbitrary rules (e.g. "you cannot do part A of a subtask followed immediately by part B of the same subtask). There are no cues as to when to switch tasks, and although a clock is present, it is covered, so that checking it has to be a deliberate action. Since 15 minutes is not long enough to complete the subtasks, subjects must remember to switch from one to another before reaching completion, so that they have at least attempted all subtasks. Thus this paradigm has a strong component of voluntary time-based task switching, which may be considered one form of prospective memory.

Despite their excellent general cognitive skills, AP and the other cases reported by Shallice and Burgess all performed these tasks below the 5% level compared with age- and IQ-matched controls. On the MET the subjects made a range of types of error. Many of these could be interpreted as problems with prospective memory. For instance they would find themselves having to go into the same shop more than once to buy items that could all have been be bought at one visit; they forgot to carry out tasks that they had previously learnt that they needed to do, or to follow task rules. They also made a range of social behaviour errors (e.g. leaving a shop without paying; offering sexual favours in lieu of payment). Shallice and Burgess (1991) rather inelegantly termed this kind of behavioural disorganization in the context of preserved intellect and other cognitive functions the "Strategy Application Disorder".

It was not possible on the basis of Shallice and Burgess's data to speculate on the anatomical localization of the lesion critical for this pattern of deficit, since the patients had suffered large traumatic lesions that invaded many subregions. Two years later however, Goldstein et al. (1993) described a case that began to suggest a possible locus. This 51-year old right-handed man (GN) had undergone a left frontal lobectomy 2.5 years earlier following the discovery of a frontal lobe tumour (mixed astrocytoma-oligodendroglioma). A 5 cm resection of left frontal lobe from the frontal pole was undertaken. This surgery made little difference to his general cognitive abilities (e.g. WAIS-R VIQ 129, PIQ 111; story recall immediate 75-90th%ile, delayed 50-70th; Rey Osterreith delayed figure recall 80-90th%ile; Trail-making 70-75th%ile). However it was nevertheless clear from his everyday behaviour that something was seriously wrong. He had held a senior management position within an international company, but two years after surgery he had to take medical retirement because of increasing lethargy. He worked from home as a freelance management consultant, but had difficulty making decisions, culminating in his taking two weeks to decide which slides to use for a work presentation, but never actually reaching a decision. He also experienced anger control difficulties.

Goldstein et al. (1993) administered Shallice & Burgess's (1991) Multiple Errands Test. GN made significantly more errors than controls, being less efficient (e.g. having to return to a shop), breaking tasks rules (e.g. using a stamp that another customer gave him), misinterpreting tasks (e.g. sticking the stamp on the wrong card), as well as failing to complete some tasks altogether, reporting that he had known he had to do them but somehow "forgot" them. He also showed some "social rule" breaks. For instance, he had omitted to find out the price of tomatoes while earlier in the greengrocers, and realizing that he should not go back into the shop unless he was to buy something, he very conspicuously climbed onto the fruit display outside the shop and peered in the shop window.

This case, and others reported in the literature, show a remarkably similar pattern of neuropsychological test performance. Burgess (2000a) summarized the performance of 8 well-known cases: None of the cases had any language or visuoperceptual impairment and all scored within the superior range on tests of current intellectual functions. Four of the seven showed no impairment on any memory test. But most remarkably, two showed no impairment on a range of clinical executive function tests known to be sensitive to frontal lobe lesions. Moreover, no executive test has been failed by more than 2/8 cases. Most remarkably, two tasks have been administered to all the patients – the Wisconsin Card Sorting Test (WCST) and Verbal Fluency – and have been performed well by every case. This contrasts with the observation that all of the reported cases of "strategy application disorder" who have been given either the Multiple Errands or Six Element Tests have failed at least one of them.

3. Isolating the critical cognitive impairment

The multitasking situations presented to a participant by the two experimental paradigms developed by Burgess & Shallice (MET and SET) share a number of similarities. These are:

1. A number of discrete and different tasks have to be completed.
2. Performance on these tasks needs to be dovetailed in order to be time-effective.
3. Due to either cognitive or physical constraints, only one task can be performed at any one time.
4. The times for returns to task are not signaled directly by the situation.
5. There is no moment-by-moment performance feedback of the sort that participants in many laboratory experiments will receive. Typically, failures are not signaled at the time they occur.
6. Unforeseen interruptions, sometimes of high priority will occasionally occur, and things will not always go as planned.
7. Tasks usually differ in terms of priority, difficulty and the length of time they will occupy.
8. People decide for themselves what constitutes adequate performance.

Burgess (2000a; Burgess et al., 2006) makes two related points about these characteristics: (a) None of them are clearly shared by traditional neuropsychological tests, nor the types of procedures that are typically used in a neurological examination; (b) They are however very common features of everyday life situations, such as Penfield and Evans' (1935) example of preparing and cooking a meal, or indeed many other situations both domestic and work-related that last more than a few minutes. So if the locus of the impairment in the frontal-lobe-based "strategy application disorder" is with processes that help deal with these situations, then the assessment problem might indeed be simply a result of the correct procedures not being typically administered in an office-based examination, rather than that there is something inherent in the office situation that suppresses the impairment, as Mesulam proposed.

4. Dispelling task, resource, and other artefacts

The single cases described above might suggest that there are some cognitive resources that underpin complex behavioural organization and multitasking that are not well measured by traditional psychometric instruments, such as IQ tests, or even tests of executive function such as the WCST, verbal fluency, Stroop, Tower of London and the like (see Burgess, 2000 for review). These results suggest that there might be executive abilities that are specific to complex behavioral organization functions like multitasking. However, before such a conclusion can be reached, there are two potential alternative explanations that have to be dismissed. First, the results might be a consequence of using tasks of differing "difficulty". In other words, that the MET and the like are somehow "more difficult" than e.g. IQ or other executive function tests, and so are better at picking up weaknesses. Second, perhaps real-life tasks like the Multiple Errands Test tend to sample behaviour over longer periods of time than many neuropsychological tests, so give more chance for the observation of failure. In order to dismiss these sorts of potential artefacts, one needs ideally to demonstrate a double dissociation between these measures (see Shallice, 1988 for an outline of the double dissociation method). In other words, one needs to demonstrate not only that there are cases who fail tests like the MET whilst passing these other tests, but also that there are those that fail these other tests but pass the MET. (See Burgess, 1997 for further data and discussion on the relation between IQ and tests of executive function.)

This pattern (normal MET in the context of poor background test performance) has never before been described, to our knowledge. However previously unpublished data from the Alderman et al. (2003) series contained two such cases. Both these cases showed normal MET performance despite marked impairments on other neuropsychological tests.

5. Cases C4 and ULI: Normal MET performance despite poor performance on other psychometric measures.

The first case, (Case C4; see Table 1) was a 50-year old male who suffered a severe brain injury in a road traffic accident. Brain scan reports described damage principally within the right occipital and left temporal regions, although given the severity of the injury it would make it difficult to rule out the possibility of damage elsewhere in the brain. His carers noticed in everyday life a tendency towards euphoric mood tendencies plus memory problems, reporting that he would get events mixed up with each other, and get confused about their correct order. He also showed some planning issues in everyday life, and on occasion these could have had severe consequences. For example, he became interested in photography, and made a spotlight by fixing a light bulb into a plastic washing up bowl, unaware of the safety implications until staff pointed this out to him.

The DEX questionnaire (part of the BADS battery) was administered to his carers, to gain an assessment of their opinion about his dysexecutive symptoms. His overall rating was 24/80, which is low average for ABI (falling at less than 38th %ile for this group), and a little above average for neurologically healthy controls (approx 77th %ile). The highest rating (3) was assigned to question 1, which asks about problems with abstract thinking. But he was additionally rated as having noticeable problems on four of the items that load upon the "Inhibition" factor of the scale (see Burgess et al, 1998 for details of DEX factor structure).

These everyday life problems were reflected in his scores on the neuropsychological tests. Although his visual memory performance (AMIPB Figure; RMT Faces) was normal, he was poor on a range of verbal recall measures (see Table 1). He also showed very poor planning skills. This was shown in particular on the Key Search and Zoo-Map tests of the BADS battery. The Key Search subtask asks participants to show how they would go about finding a key if they had lost it in a field, and aims to be a test of open-ended planning. The score is based on a number of criteria mainly relating to the efficiency of the strategy displayed, and C4 scored at the 5% level for his age. He was similarly poor at the Temporal Judgment subtask, which is an analogue of Shallice and Evans's (1978) Cognitive Estimates Test. Participants are asked to give estimates of the length of times that things task (e.g. building a house), and is thus also quite open-ended. But it was the impairment on the Zoo-Map test that was most striking. This test aims to assess the ability to formulate and implement a plan (Section 1), which is contrasted with the ability to follow a pre-formulated plan (Section 2). It involves either plotting (Section 1) or following (Section 2) a route through a map of an imaginary zoo that conforms to a set of rules. The score is based on the successful implementation of the plan, and penalties are imposed for rule breaks and lack of speed. On section 1 of the Zoo-Map test, he was at the 1%ile level for planning time and errors, and below the 5% level for performance time; overall, this section 1 performance was outside the normal range. On section 2 he was outside the normal range for errors ($N = 14$), and at the 1% level for sequencing score; overall, section 2 performance was also outside the normal range (i.e. never occurred in the normative sample of Wilson et al, 1996).

However, despite these problems with memory and planning tests he performed at the 50%ile level on the simplified version of the Multiple Errands Test (MET-SV, Alderman et al, 2003) (see Table 1). Even using the weighting system developed by Alderman et al (2003), which scores errors according to the frequency with which they appear in controls' error protocols, his performance was still average (C4's weighted error score = 7; controls mean = 4.8, SD 3.71).

In summary, this patient presented with weak verbal memory, and marked difficulties of tests of executive function that had a large planning and/or abstract reasoning and estimation component, which reflected his problems in everyday life. However despite these executive and memory problems, he was completely unimpaired on the Multiple Errands Test.

Case ULI was a 25-year-old male, employed as a warehouse worker when he sustained a very severe brain injury in a fall at work three years previous to testing. Glasgow Coma Scale (GCS) at the time of TBI was 3/15, and a CT brain scan at that stage showed bilateral frontal contusions, subarachnoid and right subdural bleeds, intracranial air and compound fractures of the right temporal bone with fractures of the left frontal and left parietal bones and of the left superior orbital wall with significant oedema. At the time of testing, carers and relatives reported a number of dysexecutive symptoms in everyday life. The most significant of these was euphoria, but also noticed regularly were problems with abstract thinking, impulsivity and disinhibition, poor planning and distractibility, and a lack of concern and insight.

His performance on most psychometric measures reflected the severity of this brain injury. His perfor-

Table 1
Examples of good Multiple Errands Test (simplified version; MET-SV) performance in the context of variable memory, IQ and executive function ability. These cases dispel the notion that the reason that dysexecutive patients fail the MET-SV is because it is "more difficult" than other neuropsychological tests, or because it samples behaviour over a longer period of time

WAIS III (age-SS)	**C4**	**ULI**	**Perception**	**C4**	**ULI**
Vocabulary	8	7	*VOSP*		
Similarities	4	9	Incomp lett.	Pass	Pass
Arithmetic	11	10	Silhouettes	Pass	Pass
Digit Span	15	8	Object des	Pass	Fail (<5%)
Information	8	9	Prog silhou	Pass	N/A
Comprehension	9	13	Dot counting	Pass	10 Pass
Letter numb seq	N/A	12	Pos Discrim	Pass	Fail (<5%)
Picture Complet	9	5	Num locat	Pass	Pass
Digit Symbol	6	2	Cube analy	Pass	Pass
Block Design	9	4			
Matrix reasoning	9	5	**Language**		
Picture arrange	10	3	*SCOLP*		
Symbol search	N/A	1			
Object Assembly	N/A	2	Speed of comprehension	> 50%ile	< 25%ile
Verbal IQ	94 Average	86 Low Average			
Performance IQ	90 Average	62 Extremely Low	Spot the word	> 10%ile	N/A
FSIQ	91 Average	73 Borderline			
Memory			**Executive**		
AMIPB Story Recall	*%ile*	*%ile*	*BADS*		
Immediate	<10	<1	Rule shift	50%ile	*<1%ile*
Delayed	<5	<1	Action program	50%ile	*25%ile*
			Key search	5%ile	*1%ile*
AMIPB Fig. Recall			Temp Judge	5%ile	*<1%ile*
Copy	>25	<5	Zoo Map		
Immediate	>50	<1	(Planning)	< 1%ile	*<1%ile*
Delayed	>50	<1	Six Elements		
			-Tasks attempted	*1%ile*	*50%ile*
AMIPB List Learning			-Max time	*1%ile*	*10%ile*
A1-A5	<5	<5	-Rule breaks	*50%ile*	*1%ile*
A6	10	<25	BADS Overall	Borderline	Impaired
B	25	<10			
			MET-SV		
Paired Associates	N/A	>25	*Task Failures*	2 normal	4 normal
			Rule Breaks	2 normal	2 normal
RMT			*Interpretations*	0 normal	0 normal
Words	SS 8	N/A	*Inefficiencies*	0 normal	0 normal
Faces	SS 11	N/A	*Overall*	50%ile	25-50%ile

mance IQ of 62 puts him in the "extremely low" range, with very poor performances on Symbol Search, Object Assembly, and Digit Symbol. On memory tests, his performance was similarly impaired. On recall of both a Story and complex figure (AMIPB) his performances were outside the normal range (age-adjusted %iles 0.8 and 0.2 respectively). His performance on tests of simple visuo-perceptual functions was normal, but he had difficulty with more complex tasks (e.g. VOSP Cube analysis SS3), and on several subtests of the BADS executive test battery his performance was outside the normal range (Rule shift, Temporal Judgment and Zoo-Map tests all 0). However, and notably, his performance on the modified Six Element Test from the BADS was normal.

Also normal was his performance on the Multiple Errands Test (MET-SV). There were two instances of rule-breaks, and he failed to achieve 4 tasks. But these scores are well within the normal range, and he made no interpretation failures or inefficiencies. His overall score of 6 errors is between the 25th and 50th percentile of healthy controls (Alderman et al., 2003).

These two cases both demonstrate that poor everyday-type multitasking performance (MET) demonstrated in the context of unimpaired performance on tests of IQ, memory etc., is unlikely to reflect artefactual influences such as "task difficulty" or length of task. If this was the case, then one would not expect to find cases who fail IQ, memory etc. tests but who perform the MET well. Moreover, one cannot dismiss the results from these two cases as being due to a lack of "sensitivity" of the MET: Not only does such an explanation run counter to the evidence from cases such as those reported by Shallice and Burgess (1991), but also group studies have shown that the MET-SV (and the Six Element Test) are *more* sensitive to general neurological involvement in this population than many other executive tasks (see e.g. Burgess et al., 1998, 2006; Alderman et al., 2003).

Instead, these cases suggest a double dissociation between memory, IQ and complex behavioral organization, at least under certain conditions. The totality of the evidence currently available is that there are brain processes which underpin multitasking and similar behaviours that are independent from those which support performance on IQ tests or tests of retrospective memory, and are even represented separately in the brain from many other executive function abilities.

6. From single cases to group studies to critical lesion location

A criticism often made of single case studies is that the data they yield might only be applicable to certain cases or situations (see Shallice, 1988 for discussion). For instance, perhaps the cases described above had an unusual developmental history, or were naturally poor at multitasking such that even a mild general cognitive decline was enough to cause a perturbation in this ability which appeared disproportionate to others. Or perhaps they were individuals for whom little demand had been made upon these skills, and so were less practiced than is typical, or adopted unusual strategies in multitasking situations. Clearly, however if no link were found at a group level between people who show multitasking failures and deficits in e.g. memory, IQ, or traditional executive test performance, then the possibility that there is a relatively discrete cognitive substrate to multitasking is given considerable weight. Further, if a correspondence is found, in an unselected sample, between this cognitive pattern and damage to a particular brain region, then it is unlikely that the cognitive findings are an artefact of selection.

Burgess et al. (2000b) investigated these questions. Sixty acute neurological patients from an unselected series of referrals (approximately three-quarters of whom were suffering from brain tumors) and 60 age- and IQ-matched healthy controls were administered a multitasking test called the Greenwich Test. This is a multitasking test that follows the principles of the Six Element Test, but where, in contrast, the majority of the variance in performance of the test comes from rule infractions rather than task-switching problems. Participants are presented with three different simple tasks and told that they have to attempt at least some of each of the tasks in 10 minutes, while following a set of rules. One of these rules relates to all subtests ("in all three tasks, completing a red item will gain you more points than completing an item of any other colour") and there are four task-specific rules (e.g. "in the tangled lines test you must not mark the paper other than to write your answers down"). The task was administered in a form that allowed consideration of the relative contributions of task rule learning and remembering, planning, plan-following and remembering one's actions to overall multitasking performance. Specifically, before participants began the test, their ability to learn the task rules (by both spontaneous and cued recall) was measured; this measure was called "Learn". They were then asked how they intended to do the test, and a measure of the complexity and appropriateness of their plans was gained (a variable called "Plan"). This enabled us to look at whether their failures could be due to poor planning (see e.g. Kliegel et al, 2005). The participants then performed the task itself and by comparing what they did with what they had planned to do, a measure of "Plan Following" was made. Multitasking performance (the number of task switches minus the number of rule breaks) was referred to as the test "Score". After these stages were finished, subjects were asked to recollect their own actions by describing in detail what they had done (variable name: "Recount"). Finally, delayed memory for the task rules was examined ("Remember").

We found that lesions in different brain regions were associated with impairment at different stages in the multitasking procedure. For instance, lesions to a large region of superior posterior medial cortex including the left posterior cingulate and forceps major gave deficits on all measures except planning. Remembering task contingencies after a delay was also affected by lesions in the region of the anterior cingulate. Critically, however, Burgess et al found that patients with left hemisphere rostral PFC lesions, when compared with patients with lesions elsewhere, showed a significant multitasking impairment (i.e. the variable "Score") despite no significant impairment on remembering task rules ("Remember" variable). Indeed, the left rostral prefrontal cases showed no significant impairment on any variable except the one reflecting multitasking performance. In other words, despite being able to learn the task rules, form a plan, remember their actions, and say what they should have done, they nevertheless did not do what they said that they intended to do.

But perhaps these results were specific to this particular multitasking test rather than being true of a class of situation with the characteristics above? This appears not to be the case. Burgess, Veitch and Costello (submitted) administered a new version of the Burgess et al. (1996) Six Element Test (SET) to sixty-nine acute neurological patients with circumscribed focal lesions and sixty healthy controls, using the administration framework of Burgess et al. (2000b). The SET differs from the Greenwich Test in that the multitasking score reflects mainly voluntary time-based switching rather than rule-following. Compared with other patients, those whose lesions involved the rostral prefrontal regions of the right hemisphere made significantly fewer voluntary task switches, attempted fewer subtasks, and spent far longer on individual subtasks. As with the study of Burgess et al (2000b), these multitasking deficits could not be attributed to deficits in general intellectual functioning, rule knowledge, planning, or retrospective memory. Furthermore, there was no obvious relation between SET performance and concomitant failure on several traditional tests of executive function.

Considering now the previous single case studies in the context of these group study findings, it is clear that there is a remarkably consistent finding of involvement of rostral PFC in cases who have high-level disorganization in everyday life. For instance, in the 6 cases reviewed by Burgess (2000a) for whom good brain scan data was available, all of them had rostral PFC involvement of either the left or right hemispheres (or both). In addition to these cases, we might also add the more recent case of Bird et al. (2004) who had suffered a rare form of stroke affecting the medial aspects of Area 10 bilaterally, and who failed the Six Element Test, despite passing some other executive tests (e.g. the WCST).

7. What is known about the functions of rostral PFC?

Given the link between dysfunction in rostral PFC and gross behavioural disorganization, it seems reasonable to assume that considering more broadly the functions of this region might shed some light upon the nature of the patients' difficulties. At the time of Mesulam's article virtually nothing was known about the functions of this region, and it was largely ignored by cognitive neuroscience researchers. This situation changed rapidly, and necessarily, with the introduction of the functional neuroimaging method. It was noticed quickly by researchers that incidental findings of rostral PFC activation were extremely common. But there seemed to be little obvious prima facie similarity between those paradigms that provoked area 10 activation. Indeed, Area 10 activations could be found during the performance of just about any kind of task, ranging from the simplest (e.g. conditioning paradigms; Blaxton et al., 1996) to highly complex tests involving memory and judgement (e.g. Burgess et al., 2001, 2003; Frith and Frith, 2003; Koechlin et al., 1999) or problem-solving (e.g. Christoff et al., 2001; for review, see Grady, 1999; Ramnani and Owen, 2004). There are now a range of hypotheses about the mental capacities that this region supports (Burgess et al., 2001, 2003; Braver et al., 2002; Bunge et al., 2005; Christoff et al., 2003, 2004; Gilbert et al., 2005; Koechlin et al., 1999, 2000, 2007; Reynolds et al., 2006; Simons et al., 2005a,b; Turner et al., 2008; for review see Burgess et al., 2005, 2006; Grady, 1999; Ramnani and Owen, 2004). There have also been a very wide range of investigations yielding results relevant to understanding the functions of this region (see Gilbert et al., 2006a), and the rate at which they are appearing is increasingly with great rapidity. At this, albeit early, stage of our understanding five broad findings have emerged from the literature (see Burgess et al., 2005, 2006, 2007 for review):

1. There is no evidence that rostral PFC lesions impair substantially performance on traditional neuropsychological measures of intelligence, routine memory functions (e.g. forced-choice recognition) or the basic substrates of language, perception or motor control.
2. But they can cause impairments with complex behavioural organisation in non-routine situations (e.g. Shallice and Burgess, 1991; Burgess, 2000a; Burgess et al., 2000b).
3. Functional neuroimaging demonstrates evidence of neural operation within this region during a wide range of tasks (Gilbert et al., 2006a).
4. In some regions of rostral PFC, this activation seems invariant according to the precise form of the stimuli being presented (e.g. whether it is pictures, words, number etc.) (Burgess et al., 2001, 2003; Gilbert et al. 2005; Simons et al., 2005a,b, 2006).

5. The activations in some regions of rostral PFC can also be remarkably indifferent to the particular mental operations being performed upon the presented stimuli, as long as these are well-rehearsed and routine, such as simple mental arithmetic or magnitude judgements, judgements of colour, reading etc. (see e.g. Burgess et al., 2001, 2003; Gilbert et al., 2005; Simons et al., 2005a,b, 2006).
6. There is evidence from neuroimaging for functional specialisation within rostral PFC (see Burgess et al., 2007; Gilbert et al., 2006a,b, 2007). In other words, different subregions of rostral PFC (area 10) are associated with neural activation during different tasks. Lateral rostral PFC regions show activation most frequently during memory tasks, with activation during tasks involving multi-task co-ordination showing a propensity for anterior rostral PFC location, whereas tasks involving mentalizing and related functions most consistently activate regions on the medial caudal surface (Gilbert et al., 2006a).

From these findings, one might suppose that rostral PFC supports, at least in part, some processing which is very "central" in character (see Shallice, 1988 for definition) (findings 3 & 4). In other words, they may not be very specifically tied to a particular modality or form of input (or output), or the precise form of the stimuli. We might also suppose that these processes are those most often used in dealing with non-routine situations (findings 1 & 2), and are quite commonly used in them (finding 2).

One theory that seeks to explain what this processing might be is termed the "gateway hypothesis" (e.g. Burgess et al., 2005, 2006, 2007a,b; see also Gilbert and others, 2005, 2006a, 2006b). It proposes that rostral PFC supports a processing system that facilitates biasing, to a novel degree, of the relative weight given to either stimulus-oriented, or stimulus-independent attending. Hence the theory rests upon a distinction between stimulus-oriented (SO) and stimulus-independent (SI) attending (McGuire and others, 1996). Stimulus-oriented attending refers to the attending behaviour that is required to detect change in the external world, via one's basic sensory systems. Examples of SO attending range from performance of vigilance tasks, to reading, watching the television, listening to a conversation and so forth. By contrast stimulus-independent attending is the attending behaviour required to concentrate on, or experience, self-generated thought ("off-line thinking"). Self-generated thought is cognition that goes beyond the overlearned associations or semantic memories provoked by currently available stimuli. Examples range therefore from task-irrelevant thoughts such as mind-wandering or daydreaming, to goal-directed cognition such as that involved in making up a novel story, or maintaining a representation over a delay period, and so forth. An everyday example to demonstrate the contrast between these modes of attending might be where one is trying to concentrate on a rather dull lecture (SO attending) versus imagining what one might do that evening after the lecture (SI attending). The gateway hypothesis proposes that rostral PFC in part supports a system that operates when one is required to maintain either mode of attending to an unusual degree, or voluntarily switch between them. More specifically, it proposes that the most anterior aspects of medial rostral PFC plays a part in supporting stimulus-oriented attending, whereas lateral rostral PFC is more involved in switching to, and maintaining, stimulus-independent cognition. In this way, the cognitive system supported by rostral PFC was characterized as a "gateway" between mental life and the external world.

The plausibility of this idea was first investigated in two neuroimaging studies by Gilbert and others (Gilbert et al. 2005, 2006c). These established that performing tasks using stimuli presented by visual display (i.e. SO attending; see Fig. 1 panel A, upper display) activates the most anterior portions of medial rostral PFC (BA 10), compared with performing the same tasks "in one's head" (i.e. SI cognition; see Fig. 1, panel A, lower display). It also showed that lateral BA 10 is activated at the point when one switches from performing a task "in one's head" to using displayed stimuli, and vice-versa (Fig. 1, panel B, which also shows the medial rostral PFC activation involved in the SO attending condition). Gilbert et al's second study replicated these findings, and showed that medial rostral PFC activations can be provoked even by simple reaction time tasks. Unpublished data from this second experiment also showed that lateral rostral PFC is activated bilaterally during periods of extended stimulus-independent cognition, and not only at the SI/SO switch-points (SI-SO contrast: left hemisphere: -40, 36, 24; BA 9/46/10; $z = 4.28$, cluster size $= 403$ voxels; right hemisphere: 38, 44, 32; BA 9/46/10; $z = 4.43$, 643 voxels; both $p < 0.001$ uncorrected (Gilbert, personal communication)).

On these grounds, one might wonder what kinds of functions, or tasks, might require the operation of this attentional "gateway". In fact there are several functions involving unusual degrees of stimulus-independent attending. Many source- or context-memory tasks for instance will make demands up-

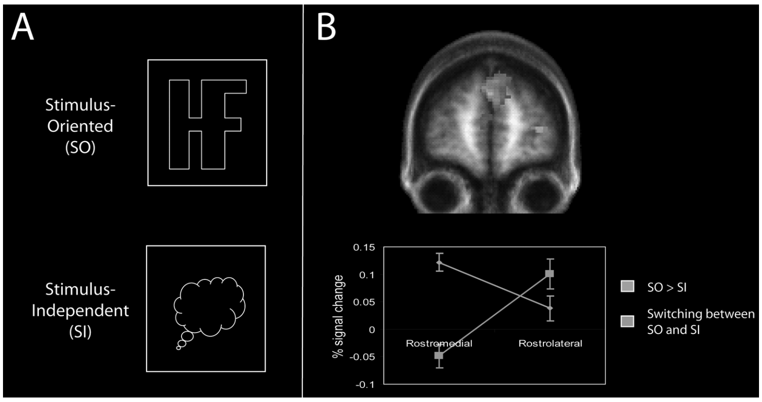

Fig. 1. Results from Gilbert et al.'s (2005) investigation of the gateway hypothesis. Panel A: one of the three tasks used. During stimulus-oriented (SO) phases of the task, participants repeatedly pressed one of two response buttons, as if navigating around the edge of a complex shape in a clockwise direction, to indicate whether the next corner would require a left or right turn. During the stimulus-independent (SI) phase, this shape was replaced by a 'thought bubble' shape and participants were required to continue navigating as before, using an internal representation of the shape presented in the SO phase. Panel B: consistent regions of activation and percent signal change in the three tasks associated with i) SO versus SI phases of the task, and ii) switching between SO and SI phases. These results indicate a dissociation between lateral and medial aspects of area 10.

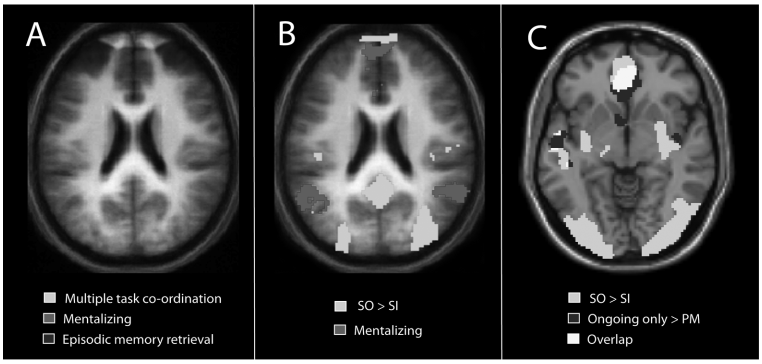

Fig. 2. Functional specialization within area 10. Panel A: regions associated with multiple task co-ordination, mentalizing, and episodic memory retrieval in a meta-analysis of 104 functional neuroimaging studies (Gilbert et al., 2006a). Panel B: distinct regions associated with i) stimulus-oriented (SO) versus stimulus-independent (SI) cognition, and ii) mentalizing in Gilbert et al. (2007). Panel C: partially overlapping regions involved in i) SO versus SI cognition, and ii) ongoing-only versus prospective memory (PM) task performance in Benoit et al. (in preparation).

on this system because of the demand for stimulus-independent attending required as a precursor to retrieval. It is axiomatic that context and/or source memory tasks require the integration of stored representations in a novel form, and/or other complex retrieval manipulations such as making novel comparisons (e.g. to determine the order of events or stimuli, in the case of memory for order (for examples see e.g. Simons et al., 2005a,b, 2006). This requires creating representations "off-line", i.e. that extend in scope far beyond the currently perceived stimulus, or its immediate associations, thus requiring stimulus-independent attending.

However one of the most obvious functions that will require the sort of attentional gateway described above is prospective memory (PM), i.e. remembering an intention over a delay period, whilst one is performing another task. Prospective memory tasks consistently provoke activation of neural populations in rostral PFC (see Burgess et al., 2008 for review). Recently, we sought to test directly the hypothesis that at least some of the neuronal populations involved in prospective memory are also involved in "attentional gateway" functions (Benoit et al., in preparation). Using a 2x2 factorial design, we crossed a prospective memory vs. ongoing task only manipulation (factor one) with a stimulus-oriented vs. –independent manipulation (factor two). We found common regions of activation within rostral PFC provoked by both manipulations, thus supporting the hypothesis. However, additional rostral PFC recruitment outside these areas suggested that this is not a complete explanation of the rostral PFC haemodynamic changes during prospective memory performance (see Fig. 2, panel C).

Another function is supported, at least in part, by rostral PFC seems to be "mentalizing". Gilbert et al.'s (2006a) meta-analysis found that medial caudal aspects of rostral PFC were most commonly active when people are performing this broad class of tasks, which might include e.g. "theory of mind" tasks, and others involving self-judgment or self-reflection (Frith and Frith, 2003, 2006; see Fig. 2, panel A). This region seemed, prima facie, to be different from those involved in the stimulus-oriented attending conditions outlined above, thus suggesting potential functional specialisation within rostral PFC. So we used fMRI to investigate medial PFC activity related to these two functions in a factorial design. This revealed adjacent but clearly distinct regions of activation related to mentalizing (relatively caudal/superior) and stimulus-oriented attending (relatively rostral/inferior), suggesting a new axis of functional organization within MPFC (Fig. 2, panel B).

8. Conclusion

Mesulam published his editorial at a time when very little was known about the functions of rostral prefrontal cortex. Had he known what we know now, the fact that some patients with frontal lobe involvement and consequent deficits in everyday life were not failing traditional cognitive tests in the clinic might have seemed less surprising. A large number of such cases, at least in the current literature, seem to have rostral PFC involvement. Although we clearly still have a great deal to learn about the functions of this large brain region, we, and others, have already shown that it supports a range of human abilities that have, until recently, received little attention from scientists or clinicians. These include those relating to complex human behavioural organisation (especially those over long time periods), meta-memory, voluntary control of attention, prospective memory, and mentalizing. No doubt many others will be discovered in time. However the critical point as regards Mesulam's mystery is that none of these functions were typically screened for in a neurological or neuropsychological investigation in 1986. Nor, indeed, is this yet common. Thus it seems likely that it is not the "office situation" here that is the constraining factor, at least in principle, but that we only have well-established tools to investigate a fraction of the full range of human abilities. Given the high incidence of rostral PFC involvement in many of the most common neurological conditions (e.g. TBI, stroke) and the impossibility of treating problems that one cannot first measure, it must surely be a priority that new tools are created to address this lacuna. In the meantime, for those abilities recently discovered to be supported by rostral PFC, such as prospective memory and multitasking, and for which standardised tests now exist, it might be wise to consider them for standard use in the assessment of executive problems, and to do so perhaps even in preference to those where a relation with impairments in everyday life is more obscure (see Burgess et al., 2006 for review).

Acknowledgments

This work was supported by Wellcome Trust grant 061171 to PWB.

References

Alderman, N., Burgess, P.W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society, 9*, 31-44.

Ackerly, S.S. & Benton, A.L. (1947). Report of a case of bilateral frontal lobe defect. *Research Publications: Association for Research in Nervous and Mental Disease, 27*, 479-504.

Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. (2005) The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends in Cognitive Sciences, 9*, 159-162.

Bird, C.M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain, 127*, 914-928.

Blaxton, T.A., Zeffiro, T.A., Gabrieli, J.D.E., Bookheimer, S.Y., Carrillo, M.C., Theodore, W.H. & Disterhoft, J.F. (1996). Functional mapping of human learning: A positron emission tomography activation study of eyeblink conditioning. *J of Neurosci, 16*, 4032-4040.

Brandimonte, M., Einstein, G.O., & McDaniel, M.A., (1996). Prospective Memory: Theory and Applications. Lawrence Erlbaum Associates, Mahaw, New Jersey.

Braver, T.S. and Bongiolatti, S.R. (2002). The role of the frontopolar prefrontal cortex in subgoal processing during working memory. *Neuroimage, 15*, 523-536.

Brickner, R.M. (1936). The intellectual functions of the frontal lobes: A study based upon observation of a man after partial bilateral frontal lobectomy. New York: Macmillan.

Bunge, S.A., Wendelken, C., Badre, D., & Wagner, A.D. (2005). Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. *Cereb Cortex, 15*(3), 239-249.

Burgess, P.W., Alderman, N., Emslie, H., Evans, J.J., Wilson, B.A., & Shallice, T. (1996). The simplified six element test. In Behavioural Assessment of the Dysexecutive Syndrome (eds. B. A. Wilson, N. Alderman, P. W. Burgess, H. Emslie and J. J. Evans), Bury St. Edmunds, U.K.: Thames Valley Test Company.

Burgess, P. W., Alderman, N., Evans, J., Emslie, H. & Wilson, B. A. (1998) The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society 4*, 547-558.

Burgess, P. W. (1997) Theory and methodology in executive function research. In: P. Rabbitt (Ed.) Theory and Methodology of Frontal and Executive Function (pp. 81-116). Hove, U.K.: Psychology Press.

Burgess, P.W. (2000a). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research, 63*, 279-288.

Burgess, P.W., Veitch, E., Costello, A. & Shallice, T. (2000b). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia, 38*, 848-863.

Burgess, P.W., Quayle, A. & Frith, C.D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia, 39*, 545-555.

Burgess, P.W., Scott, S.K. & Frith, C.D. (2003). The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. *Neuropsychologia, 41*, 906-918.

Burgess, P.W., Simons, J.S., Dumontheil, I., & Gilbert, S.J. (2005). The gateway hypothesis of rostral PFC function. In: J. Duncan, L. Phillips and P. McLeod (Eds.) Measuring the Mind: Speed, Control and Age (pp. 215-246). Oxford University Press.

Burgess, P. W., Gilbert, S. J., Okuda, J. & Simons, J.S. (2006) Rostral Prefrontal Brain Regions (Area 10): A Gateway between Inner Thought and the External World? In: W. Prinz & N. Sebanz (Eds.), Disorders of Volition (pp. 373-396). Cambridge, MA: MIT Press.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M-A., Dawson, D. R., Anderson, N. D., Gilbert, S. J., Dumontheil, I. & Channon, S. (2006) The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society 12*, 1-16.

Burgess, P.W., Gilbert S.J., & Dumontheil, I. (2007). Function and localization within rostral prefrontal cortex (area 10). *Philosophical Transactions of the Royal Society of London B, 362*, 887-899.

Burgess, P.W., Dumontheil, I., & Gilbert S.J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn Sci, 11*(7), 290-298.

Burgess, P.W., Veitch, E.J. & Costello, A. (Submitted). Multitasking deficits in humans following rostral prefrontal lesions: The Six Element Test.

Burgess, P.W., Dumontheil, I., Gilbert,S.J., Okuda,J., Schölvinck, M.L., & Simons, J.S. (2008). On the role of rostral prefrontal cortex (area 10) in prospective memory. in Kliegel, M., McDaniel, M.A., and Einstein,G.O. (eds), Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives. Mahwah: Erlbaum.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J.K., Holyoak, K.J., & Gabrieli, J.D.E. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage, 14*, 1136-1149.

Christoff, K., Ream, J.M., Geddes, L.P., & Gabrieli, J.D. (2003). Evaluating self-generated information: Anterior prefrontal contributions to human cognition. *Behav Neurosci, 117*, 1161-1168.

Christoff, K., Ream, J.M., & Gabrieli, J.D.E. (2004). Neural basis of spontaneous thought processes. *Cortex, 40*, 1-9.

Eslinger, P.J., & Damasio, A.R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient E.V.R. *Neurology, 35*, 1731-1741.

Frith, U. & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Phil Trans R Soc B, 358*(1431), 459-473.

Frith, C.D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron 18*, 50(4), 531-534.

Gilbert, S.J., Frith, C.D. & Burgess, P.W. (2005). Involvement of rostral prefrontal cortex in selection between stimulus-oriented and stimulus-independent thought. *Euro J Neurosci, 21*, 1423-1431.

Gilbert, S.J., Spengler, S., Simons, J. S., Steele, J.D., Lawrie, S.M., Frith, C.D. & Burgess, P.W. (2006a). Functional specialisation within rostral prefrontal cortex (area 10): A meta-analysis. *J Cogn Neurosci, 18*(6), 932-948.

Gilbert, S.J., Spengler, S., Simons, J.S., Frith, C.D. & Burgess, P.W. (2006b). Differential functions of lateral and medial rostral prefrontal cortex (area 10) revealed by brain-behavior correlations. *Cerebral Cortex, 16*(12), 1783-1989.

Gilbert, S.J., Simons, J.S., Frith, C.D., & Burgess, P.W. (2006c). Performance-related activity in medial rostral prefrontal cortex (Area 10) during low demand tasks. *J Exp Psychol: Hum Percep Perfor, 32*, 45-58.

Gilbert, S.J., Williamson, I.D.M., Dumontheil, I., Simons, J.S., Frith, C.D., & Burgess, P.W. (2007). Distinct regions of medial rostral prefrontal cortex supporting social and non-social functions. *Soc Cog Affect Neurosci, 2*, 217-226.

Goldstein, L.H., Bernard, S., Fenwick, P.B., Burgess, P.W., & McNeil, J. (1993). Unilateral frontal lobectomy can produce strategy application disorder. *J Neurol Neurosurg Psychiat, 56*, 274-276.

Grady, C.L. (1999). Neuroimaging and activation of the frontal lobes. In B. L. Miller and J. L. Cummings (Eds), The human frontal lobes: Function and disorders (pp. 196-230). New York: Guilford Press.

Kliegel, M., Phillips, L.H., Lemke, U., & Kopp, U.A. (2005). Planning and realisation of complex intentions in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry, 76*(11), 1501-1505.

Kliegel, M., Jäger, T., & Phillips, L.H. (2008). Adult age differences in event-based prospective memory: a meta-analysis on the role of focal versus nonfocal cues. *Psychol Aging, 23*(1), 203-208.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature, 399*, 148-151.

Koechlin, E., Corrado, G., Pietrini, P., & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in planning. *Proc Nat Acad Sci USA, 97*, 7651-7656.

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*(5850), 594-598.

McGuire, P.K., Paulesu, E., Frackowiak, R.S. J. & Frith, C.D. (1996). Brain activity during stimulus independent thought. *NeuroReport, 7*, 2095-2099.

Metzler, C., & Parkin, A.J. (2000). Reversed negative priming following frontal lobe lesions. *Neuropsychologia, 38*(4), 363-379.

Mesulam, M.M. (1986). Frontal cortex and behavior. *Ann Neurol, 19*(4), 320-325.

Penfield, W., & Evans, J. (1935). The frontal lobe in man: a clinical study of maximum removals, *Brain, 58*, 115-133.

Ramnani, N. & Owen, A.M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience, 5*, 184-194.

Reynolds, J.R., McDermott, K.B., & Braver, T.S. (2006). A direct comparison of anterior prefrontal cortex involvement in episodic retrieval and integration. *Cereb Cortex, 16*, 519-528.

Shallice, T., & Evans, M.E. (1978) The involvement of the frontal lobes in cognitive estimation. Cortex, 14, 294-303. (Reprinted in J. Davidoff (Ed.), Brain and behaviour: Critical concepts in psychology, Vol. 4, London: Routledge).

Shallice, T. (1988). From Neuropsychology to Mental Structure. New York: Cambridge University Press.

Shallice, T. & Burgess, P.W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*, 727-741.

Simons, J.S., Gilbert, S.J., Owen, A.M., Fletcher, P.C., & Burgess, P.W. (2005a). Distinct roles for lateral and medial anterior prefrontal cortex in contextual recollection. *J Neurophys, 94*, 813-820.

Simons, J.S., Owen, A.M., Fletcher, P.C., & Burgess, P.W. (2005b). Anterior prefrontal cortex and the recollection of contextual information. *Neuropsychologia, 43*, 1774-1783.

Simons, J.S., Schölvinck, M., Gilbert, S.J., Frith, C.D., & Burgess, P.W. (2006). Differential components of prospective memory? Evidence from fMRI. *Neuropsychologia, 44*, 1388-1397.

Turner, M.S., Simons, J.S., Gilbert, S.J., Frith, C.D., & Burgess, P.W. (2008). Distinct roles for lateral and medial rostral prefrontal cortex in source monitoring of perceived and imagined events. *Neuropsychologia, 46*, 1442-53.

Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H. & Evans, J.J. (1996). Behavioural assessment of the dysexecutive syndrome. London: Harcourt Assessment.