Language and Cognitive Functions in a Neurological Tumour Population: A Long-Term Follow-up Study

by

Katherine Fowler

A thesis submitted to Victoria University of
Wellington in fulfilment of the
Requirements for the degree of
Doctor of Philosophy in Psychology

Victoria University of Wellington

2018

Abstract

Complex cognitive capacities such as language and "executive function" are difficult to evaluate in neuropsychological populations due to their multifactorial nature. The current study takes a cognitively-motivated core-skills approach to their assessment. Across four investigations, these various capacities are decomposed into simpler core skills based on current cognitive theory. An undifferentiated sample of 28 neurological tumour patients is then assessed on these skills.

In the first study, we assessed the "core skills" underpinning language function at three time points: pre-operative (one day prior to surgery), post-operative (within three days of surgery), and at long-term follow-up (at least three months post-surgery). This approach was sensitive at detecting impairment; indeed, almost half of the patients showed persistent long-term language deficits even at long-term follow-up. The decompositional approach also proved effective at predicting long-term outcomes. Overall, these results suggest that the subtle language deficits may be more common and more persistent than previously estimated in tumour populations.

 The second study examined the relationship between "core" language skills and sentence-level language processing at long-term follow up. Whilst there were few significant correlations, the results nonetheless suggest that "core skills" measures may be useful predictors of some aspects of sentence-level processing.

The third study isolated and identified "core" skills that are essential for complex cognitive control more generally, and assessed these in our patient sample at long-term follow-up. Results were broadly supportive of this decompositional approach, and again, our assessments proved highly sensitive at detecting deficits in this patient sample.

The fourth study examined the relationship between language processing and complex cognitive control. Specifically, we examined whether there are systems specially dedicated to the control of language, or whether control functions operate across all domains. Overall, our results were broadly consistent with the domain-specific view - that there may be functionally distinct control systems operating on verbal and nonverbal material.

The results, taken together suggest that a core skills approach to neuropsychological assessment has considerable promise, and is worth exploring further in a large patient sample. This approach may also help extend our understanding of the functional organisation of language, and the broader cognitive skills necessary for linguistic operations.

Acknowledgements

It was truly a privilege to be involved in this project, and there are a number of people who made this possible. First, I would like to thank each of the patients and their families for their generous time and willingness to be involved in this project. I am enormously grateful and humbled that they allowed me to be a small part of their journey during what was an extremely difficult time in their lives. Thank you for being so open. I have learnt an enormous amount from working these patients, which I will carry throughout my personal and clinical life.

To my primary supervisor, Dr Carolyn Wilshire. I simply cannot thank you enough for your endless guidance, expertise, support, and humour throughout this project. It has truly been a privilege to work with you and learn from you. Thank you for always encouraging me and having faith in me. I would also like to thank Mr. Andrew Parker for his involvement in this project. Also, to Josh Faulkner, whose research formed the basis for this project – thank you for your constant support and guidance, and your genuine interest in my research.

To my lab mates and those in the clinical programme, particularly Emma Ashcroft, Samuel Clack, Victoria Lee, and Kris Neilson, thank you for your unwavering support, encouragement and humour throughout this journey. Also, to my friends, particularly Kate Norman and Catherine Pihema, thank you for listening and supporting me, providing me with endless cups of coffee when I needed it, and picking me up when times were tough.

To my parents, Jennie and Ray, thank you for your unconditional support, encouragement, and constant reassurance when I was convinced I couldn't do it. Thank you for listening to me, proofreading my thesis, and always encouraging me to follow my passion, no matter how hard the journey. Without you, this would have been much harder.

Finally, to my partner Jake who has been on this journey with me from the beginning. I simply cannot thank you enough for your unwavering love and support. Thank you for keeping the home fires burning, sharing the burden with me, celebrating the joy, and encouraging me to follow my dreams. You have been my biggest cheerleader and support and I truly couldn't have done this without you by my side.

This thesis is dedicated to the patients and their families who sadly lost their battle – each of whom selflessly gave up their precious time in the hope that it will help others.

I hope that can be fulfilled.

Table of Contents

Table of Tables

Table of Figures

General Overview

The ability to communicate through spoken language is crucial for everyday social functioning. However, language capabilities can be severely compromised following brain damage or neurological illness. Improving our understanding of the nature of these difficulties can ultimately lead to new insights into how best to rehabilitate affected individuals. An improved understanding of language functioning can also shed light on the nature of human language more generally, for example, the way in which healthy individuals understand and produce language, and the brain structures critical to these processes.

Whilst many studies have investigated language difficulties following stroke, fewer have examined language outcomes in other aetiologies, such as neurological tumours (Davie, Hutcheson, Barringer, Weinberg, & Lewin, 2009). From a clinical perspective, such studies are urgently needed. Given the importance of language to everyday social functioning, the ability to maximise language capacity is an important goal for both pre-surgical planning and for post-surgical interventions (Davie et al., 2009; Miceli, Capasso, Monti, Santini, & Talacchi, 2012). Also, from a more theoretical perspective, investigating aspects of language functioning in neurological tumour populations can contribute a unique source of knowledge, which can ultimately extend our understanding of the functional organisation of language, and the associative cognitive skills necessary for linguistic operations. Moreover, given the distinct localisation patterns of neurological tumours, this population offers a unique opportunity to examine the mechanisms underpinning language functions, and the cognitive control processes that may be involved in, and necessary for language functioning. Such an improved understanding can ultimately contribute newfound knowledge regarding distinct functional localisation patterns, and thus can complement our existing knowledge base of language processes in stroke populations.

It is these considerations that led to the development of the current investigation, wherein the overarching goal is threefold: a) to explore long-term language outcomes in a neurological tumour population; b) to gain new insights into the cognitive processes that underpin language dysfunction at both the single word and the sentence level; and c) to explore more general relationships between language and other aspects of cognitive function.

The thesis is separated into four investigations, each aimed at addressing distinct clinical and theoretical objectives. The objectives for each investigation are as follows:

1. To understand the long-term clinical impact of linguistic functioning in individuals who have recently undergone tumour resection surgery (Chapter 2);

2. To explore the extent to which sentence-level language can be decomposed into smaller core cognitive language components (Chapter 3);

3. To characterise and define crucial components of cognitive control in order to determine the functional and anatomical organisation of control processes (Chapter 4); and

4. To explore the relationship between language function and other types of cognitive processes that operate outside the verbal domain (Chapter 5).

To this end, we present data from a sample of 28^1 patients with primary neurological tumours. Each is assessed on a newly developed, cognitively-motivated language assessment battery at three time periods: *preoperatively* (the day prior to surgery), *postoperatively* ($1 - 3$ days following surgery), and at least *three-months post-surgery* (average time since surgery 5.7 months, $3 - 12.2$ months). During the follow up testing phase, we also assess a number of broader cognitive processes, particularly those associated with frontal lobe damage (e.g., *inhibitory control*, *internally-driven response selection, initiation/activation (speed), sustained attention,* and *performance monitoring)*. Particular emphasis will be on the inter-relationships between these processes and whether these types of high-level control functions are specific to the verbal domain, or instead are subserved by more generalised higher-level control processes that operate across the verbal and non-verbal domain equally.

---

[1] The fist investigation was limited to a subset of 25 of these patients.

Chapter 1: Overview of Language and Cognitive Theories

Recent advances in cognitive theory offer a fine-grained theoretical framework of language function, which may prove effective in assessing language in tumour patients. Specifically, these current theories propose that a number of core, distinguishable processes are involved in language functioning, each of which represent a unique cognitive process or operation.

For example, it has been suggested that lexical retrieval for spoken word production involves two distinguishable processing stages, namely, *lexical selection* or the selection of the appropriate lexical elements, and *phonological encoding*, or the subsequent retrieval of their corresponding sound information (see Figure 1.1). Of course, in most instances, a speaker cannot select the appropriate lexical item until they have retrieved a well specified semantic representation of the concept to be described (what we will refer to as *accessing semantic knowledge*) (Friedmann, Biran, & Dotan, 2013). Moreover, once phonological encoding is complete, further processing is necessary: an articulatory-motor plan for the word must be generated. If we include these additional processes, we are able to identify four distinct, incremental processing stages that are required for the production of a single word: *accessing semantic knowledge, lexical selection, phonological encoding,* and *articulatory-motor planning.* It has been argued that each of these processes can become selectively impaired after brain damage, and impairment to each is associated with a distinct neuropsychological profile (see discussion below). Before we begin this discussion, we acknowledge that within these models, there is much disagreement about the cognitive mechanisms that underpin each skill. Therefore, we present the skills for which there is widespread agreement for their existence within the literature.

Semantic description of item
to be produced

Stage 1:
Lexical selection

Word that best matches
semantic description
(sound form not yet specified)

"cat"

Stage 2:
Phonological retrieval

Phonological representation of
word

/kæt/

Articulatory-Motor
Programming
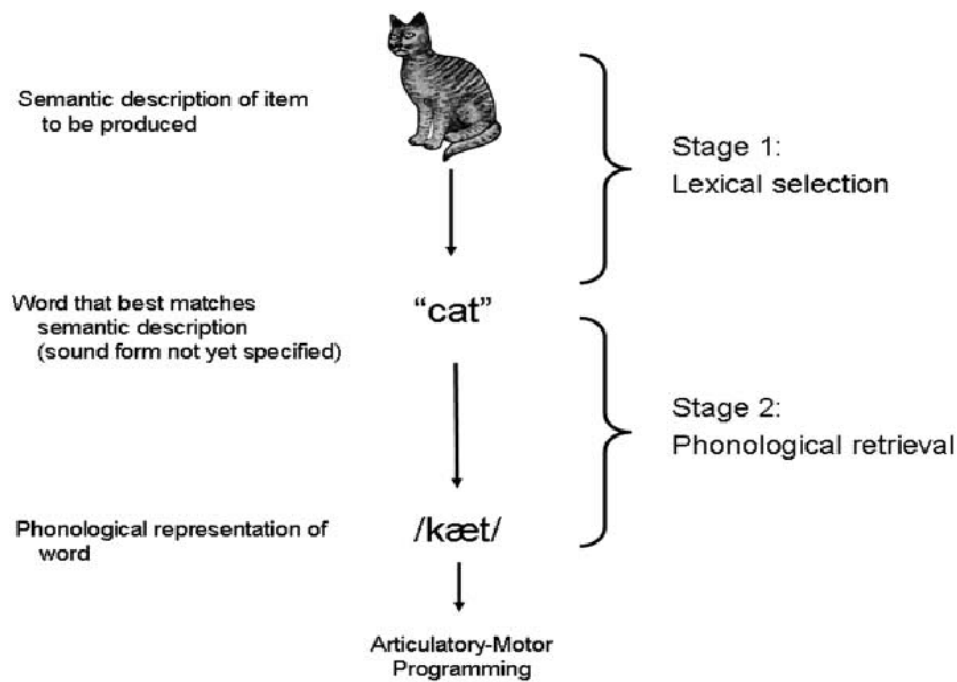
Figure 1.1. The four cognitive skills involved in single word production, with picture naming as a framework (adapted from Wilshire, 2014).

Many theories express the stages of word production within a spreading activation framework (e.g. Dell, 1986; Harley, 1984; Levelt, 1999; Roelofs, 1992; 2004). An example of such a framework is illustrated in Figure 1.2. According to Dell's (1986) model, the first stage of word production, *accessing semantic knowledge*, involves activating the units representing the semantic features of the desired concept. In this example, the word '*dog*' may activate a number of distinct attributes, such as 'bark', 'run', and 'woof'. Once elements of the semantic representation have been activated, activation then spreads to other interconnected units representing the various lexical items that are associated with these semantic features, such as 'cat', 'dog' and so on. Importantly, activation spreads in a graded manner (for example, the lexical unit representing *dog* will receive greater activation than that for *cat*, because it shares more of the target semantic features). The second stage of processing, *lexical selection*, is complete when the lexical unit that is most highly activated is selected for further processing (in this case, *dog*). An important feature of this aspect of the model is that if the target lexical item does not become sufficiently activated, for whatever reason, the next most highly activated item is likely to be one that shares some of its core semantic features. The result in this situation will be a semantic substitution error (e.g., producing *cat* instead of dog).
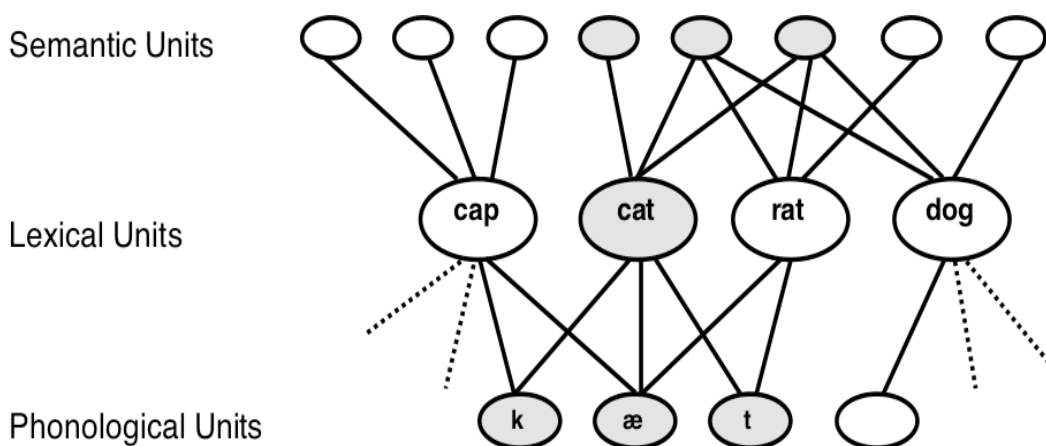
Figure 1.2. A conceptualisation of the spreading activation framework of single-word production, based upon the theory of Dell and O'Seaghdha (1991, 1992).

Within this spreading activation framework, the process of phonological encoding involves the spread of activation from lexical units to their subordinate phonological units. If the process operates effectively, the phonological units that receive the greatest activation will be those corresponding to the phoneme in the chosen word (see especially Dell, 1986; Levelt, 1999). These frameworks also propose that additional processes occur after phonological encoding, which include creating a motor programme for articulation. However, these are not usually fleshed out in spreading activation terms.

This type of framework, especially when expressed in interaction terms, leads to direct predictions as to the patterns we might expect to see if a particular processing stage was impaired. An impairment at the stage of accessing semantic knowledge would be predicted to lead to semantic substitution errors. If this process were inaccurate or incomplete, then the item that is ultimately selected may be one that shares many – but not all - of the target's semantic features. For example, the word "cat" may be produced in place of "dog" as both share similar semantic features (e.g. both are animals and household pets, have four legs and a tail). These types of semantic substitution errors have been observed in the degenerative disorder of semantic dementia, a condition largely characterised by difficulties producing and understanding words, with considerable confusion between semantically related items in both production and comprehension tasks (e.g. "*cat*" vs. "*dog*") (Jefferies & Lambon Ralph, 2006; Thompson, Patterson, & Hodges, 2003). It has been suggested that this condition reflects a difficulty accessing semantic representations. Indeed, the fact that errors

occur in both production and comprehension has been taken as evidence that the disorder involves this stage of processing (rather than the later stages that are specific to word production). The features of degeneration in semantic dementia are most commonly focused on the left anterior temporal pole (Hodges & Patterson, 2007; Pijnenburg, Gillissen, Jonker, Scheltens, 2004; Thompson et al., 2003), with damage to this region characterised by the following patterns of performance: 1) semantic errors in picture naming; 2) disproportionately poorer performance on a category, compared to letter fluency task; and 3) difficulty correctly matching pictures to a word when the distractor items are semantically related to the target (see Corbett, Jefferies, Ehsan, & Lambon Ralph, 2009; Garrard & Hodges, 2000: Garrard, Perry, & Hodges, 1997; Jefferies & Lambon Ralph, 2006).

An impairment involving the lexical selection stage is also predicted to result in semantic substitution errors. This is because where activation spreads from the semantic units to the corresponding lexical units, lexical items that share semantic overlap with the target will also become activated (Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999; Mahon, Cost, Peterson, Vargas, & Caramazza, 2007). Consequently, lexical items that are concurrently activated will "compete" for selection. If the process goes awry, it is highly probable that the item selected in place of the target will be one that shares many of its semantic features. However, in contrast to an impairment in accessing semantic knowledge, this problem would be predicted to affect only word production; comprehension of words would not be affected. Further, many spreading activation models propose that low frequency lexical items (e.g. tambourine) will have higher activation thresholds than high frequency words that are commonly retrieved (e.g. couch), therefore, they require less activation for selection. Consequently, there is a greater likelihood of making a lexical substitution error on a lower frequency than on a higher-frequency word (see Nickels, 2002). It has been proposed that the lexical selection process is selectively impaired in the aphasic syndrome of anomia (or more specifically, what is known as *classical* or *pure* anomia: Andreetta, Cantagallo & Marini, 2012; Wilshire, Keall, Stuart, & O'Donnell, 2007). This disorder has been localised to the mid to posterior portion of the left middle temporal gyrus, and is typically characterised by intact general semantic knowledge and word comprehension, yet selective word naming difficulties, with low frequency words particularly problematic. Specifically, patients typically produce a mixed pattern of errors, including semantic substitutions (e.g. dog -> cat) and and/or difficulty retrieving items from the mental lexicon, resulting in a complete omission of responses (e.g. "it has four legs and woofs, but I cannot think of the name"; Dell,

Lawler, Harris, & Gordon, 2004; Lambon Ralph, Sage & Roberts, 2000; Miceli, Amitrano, Capasso & Caramazza, 1996).

Further, within spreading activation frameworks, difficulty during the phonological encoding stage would be predicted to lead to phonological substitution errors at both the single and sentence-level. For example, an individual may say "pobacco" instead of "tobacco" (see Buchsbaum et al., 2011; Faulkner, 2015). Such errors would be predicted to be more likely under conditions that place high demands on phonological elements, such as producing words that contain bi and multi-syllabic demands (e.g. *toast-er* and *hos-pit-tal*), relative to those with mono syllables (e.g. *thanks*; Pate, Saffran, & Martin, 1987; Wilshire, 2002; Wilshire & McCarthy, 1996). This word-length effect is also predicted to occur because additional phonemes, such as those in multi-syllabic words, need to be retrieved and inserted into the correct phonological frame. Moreover, these spreading activation models propose that the phonological encoding stage is also essential for accurate repetition of aurally presented words. Consequently, they predict that an impairment at this stage of processing would also affect the capacity to repeat single-words, particularly longer words, which require additional phonemes to be retrieved and/or inserted (Caramazza, Basil, Koller & Berndt, 1981; Dell, Schwartz, Martin, & Saffran 1997). Further, it has been suggested that the phonological processing phase may be selectively impaired in *Conduction aphasia,* a disorder characterised by the production of phonological errors on all types of word production tasks (including repetition), but intact comprehension abilities (Buchsbaum, et al., 2011; Pradat-Diehl et al., 2001; Rapp & Goldrick, 2000). It is proposed that left superior temporal and inferior parietal regions are involved in phonological encoding processes. For example, in a recent voxel-based lesion symptom mapping study (VLSM) on a sample of 106 patients with post-stroke left-hemispheric lesions, phonological errors (phonemic paraphasias) on the Philadelphia naming test were associated with the left post-central gyrus, inferior precentral gyrus, and supramarginal gyrus (Schwartz, Faseyitan, Kim, & Coslett, 2012).

The final stage emphasised in current models of word production is *articulatory motor programming.* At this stage, an articulatory motor plan must be formulated and programmed for the relevant word to be correctly executed and expressed (Indefrey & Levelt, 2004; see esp. Romani & Galluzzi, 2005; Romani, Olson, Semenza, & Granà, 2002). There are a number of theories as to how this might occur. One view is that, once a syllable-sized phonological sequence has been identified, it then activates its respective entry in a

'syllabary', which contains information about the appropriate articulatory gestures required to produce that syllable (Indefrey & Levelt, 2004). Articulatory motor programming is proposed to be selectively impaired in apraxia of speech, a disorder characterised by particular difficulty with the motoric aspects of speech (Frey, Woods, Knight, Scabini, & Clayworth, 1987; Johns & Darley, 1970) as evidenced by slower speech rate, and articulation errors involving the timing and coordination of speech (Canter, Trost & Burns, 1985; Dronkers, 1996). It has been proposed that left anterior regions are involved in articulatory processes. For example, in a recent VLSM study involving 102 left hemisphere stroke patients, the left frontal gyrus and anterior temporal regions were associated with performance on the articulation and prosody subtests of the Aachen Aphasia Test (Henseler, Regenbrecht & Obrig, 2014).

Beyond Single Word Production

Clearly, there is more to language performance than just producing the names of words. Recent research in psycholinguistics has identified a number of additional processes that become particularly important when we move beyond simple naming tasks.

At the simplest level, there is clear evidence that a unique and distinct set of processes are required for the auditory identification of words. Cognitive theories of language propose that the ability to identify auditory words involves recognising the acoustic structures of a word, and subsequently mapping these onto the corresponding lexical representation (Hickok & Poeppel, 2007; Marslen-Wilson, 1987) (see Chapter three for a more detailed discussion of this process). Evidence to suggest that this is a distinct process comes from a number of studies that report the left posterior-temporal parietal regions are sensitive to the ability to identify auditory words. Specifically, patients with lesions to this region may be selectively impaired in both repeating and discriminating between real words and non-words; but may show little evidence of difficulties on spoken language production tasks. These patterns of profound deficits in auditory word recognition are often associated with Wernicke's aphasia and pure word deafness (Caplan, Gow & Makris, 1995).

Moving now to the production of utterances beyond simple naming, one capacity that is crucial for the successful production of well-formed sentences is the capacity to retrieve verbs. The act of retrieving verbs from the mental lexicon is considered a necessary and distinct process in the formulation of a sentence plan (Marshall, Pring & Chiat, 1998; see Chapter 3). Specifically, the view is held that verbs place considerable constraints on the

overall sentence structure, whereby they guide important elements of a narrative, such as specifying the argument structures and semantic nature of a sentence frame, and the roles they play in determining its meaning (e.g. Ahrens, 2003; Levelt 1989; 1999; Pickering & Branigan, 1998; Shapiro & Levine, 1990; Berndt, Haendiges, Mitchum, & Sandson, 1997; Trueswell & Kim, 1998). The left inferior frontal regions have been implicated in the retrieval of verbs; specifically, individuals with lesions to these regions exhibit disproportionate difficulty on action-naming compared to object naming tasks (e.g. Piras & Marangolo, 2007; 2010).

Recently, it has also been proposed that a specialised top-down control system operates within the lexicon to resolve conflict and competing representations (see Faulkner, 2015; Hamilton & Martin, 2005; Novick, Trueswell, & Thompson-Schill, 2005; Robinson, Shallice, Bozzali & Cipolotti, 2010; Scott & Wilshire, 2010). This system is likely to be heavily engaged in demanding speech situations, such as producing complex utterances, or generating speech where there is a high degree of competing, conflicting information. At the single-word level, a consistent finding is that damage to the left inferior frontal gyrus (LIFG) is associated with a characteristic pattern of performance under conditions that appear to require the resolution of conflict between two or more competing alternatives (see Novick et al., 2005; Novick, Trueswell, & Thompson-Schill, 2010). Some researchers have studied these processes using the verb generation task, which requires participants to view an object or read/hear its name, and then generate an action that is associated with that object. Within this task, it is possible to compare performance on nouns that are associated with a single dominant action (e.g., high word-strength words: scissors –> "*cut*"), with those that offer a number of possible alternatives and are thus low-strength words (e.g., the verb "*rope*" could be associated with '*hang, tie, pull,* and *swing*'). Some individuals with damage to the LIFG are disproportionately slow and/or inaccurate on the second type of items, when compared to the first (Cameron-Jones, 2008; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thompson-Schill, D'Esposito, & Kan, 1999). Another task that highlights specific types of difficulties in patients with LIFG damage is the Stroop task. This paradigm involves the presentation of a coloured word, with participants required to name the colour the word is written in and ignore what the word actually says. On critical items, the word is incongruent with the colour it is written in (e.g., the word 'black' written in brown ink); this situation would appear to involve a high degree of conflict resolution, and more specifically, resolution in favour of the less frequent and less well learned alterative (that is, the colour name *brown*

and not the word name, *black*). Indeed, a number of studies report that individuals with lesions to the LIFG exhibit reduced accuracy and slower latencies on these incongruent items, when compared to simple naming of colour patches, and/or naming of words where the colour name and the ink colour disagree (Hamilton & Martin, 2005; Scott & Wilshire, 2010).

Another skill that has been argued to be essential in sentence comprehension particularly is verbal short-term memory. We discuss this process in greater detail in Chapter three but will briefly describe the concept here. Within Baddeley, Lewis, and Vallar's classical conceptualisation of working memory (1984), verbal short-term memory is conceptualised as a domain-specific storage system that maintains verbal information in a phonological form for short periods (the *phonological store*). This framework also proposes the existence of an additional process, called the phonological loop, which can be used to rehearse and refresh material in the phonological store, thus enabling it to be retained for longer (for example, via strategies such as rehearsal; for a more detailed discussion of this model see Chapter three). Classically, the capacity of the phonological store – and its associated systems, such as the phonological loop - is measured using the digit span task: the participant hears a sequence of digits and must repeat them back immediately in the same order they heard them. Another task that is considered to place considerable demands on verbal short-term memory is non-word repetition. Specifically, these unfamiliar phonological items are not represented in our mental lexicon, and thus are highly dependent on a temporary phonological storage site (Archibald & Gathercole, 2007). Indeed, a number of lesion and neuroimaging studies have reported high correlations between non-word repetition and digit-span (Gathercole & Baddeley, 1989; Gathercole, Willis, Emslie, & Baddeley, 1992; Gupta, 2003; Gupta, MacWhinney, Feldman & Sacco, 2003).

Theories of Cognitive Control

So far, the emphasis of this Chapter has focused on current cognitive theories of language specifically. However, to fully understand the mechanisms that underpin complex language operations, it is also necessary to investigate higher-level, non-linguistic functions, such as cognitive control. Cognitive control is a general term used to refer to a collection of cognitive skills that are necessary to internally guide goal-oriented behaviours (Badre, 2008; Badre, Hoffman, Cooney, & D'Esposito, 2009). Being able to act in accordance with an internal goal requires a number of capacities, including the ability to formulate and maintain that goal in mind, the ability to enhance goal-relevant information and suppress irrelevant information, and also to monitor one's progress against the current goal (Hamilton & Martin,

2005; Folk, Remington, & Johnston, 1992; Ridderinkhof, Van den Wildenberg, Segalowitz, & Carter, 2004; Stuss, Shallice, Alexander, Picton, 1995; Stuss et al., 2005). Whilst the prefrontal cortex (PFC) is widely accepted as playing a key role in these functions (MacDonald, Cohen, Stenger, & Carter, 2000; MacPherson, Turner, Bozzali, Cipolotti, & Shallice, 2010; Miller & Cohen, 2001; Rueckert & Grafman, 1996; Stuss et al., 1995; 2005; Stuss, 2011), there is ongoing debate regarding the nature of the processes involved. Specifically, it continues to remain unclear how these processes are recruited in different tasks, as well as their precise neural localisations (Ridderinkhof et al., 2004, Tsuchida & Fellows, 2013). The following section briefly reviews the literature relevant to understanding cognitive control.

Cognitive Control and the Prefrontal Cortex

The ability to control higher-level cognitive processes in order to achieve an important goal has been associated most strongly with the prefrontal cortex (PFC) (Miller & Cohen, 2001; Ridderinkhof et al., 2004; Stuss et al., 1995; Stuss, 2011). Indeed, the PFC is in an ideal position to support these functions, given it projects to almost all regions of the parietal and temporal cortices, with connections to the pre-striate areas of the occipital lobe, and subcortical connections to structures including the limbic regions, basal ganglia, cerebellum and various brainstem nuclei via thalamic connections (Barbas, 1995; Ridderinkhof et al., 2004). Further, vast networks also exist within subregions of the frontal cortex, with the anterior cingulate cortex, an area heavily implicated in error detection and monitoring, projecting to almost all regions of the frontal lobe (see Ridderinkhof et al., 2004). One hypothesis is that the PFC operates by guiding the flow of biasing signals to these regions (Miller & Cohen, 2001).

Nonetheless, understanding the functional and anatomical architecture of cognitive control is challenging. Indeed, fMRI studies show that many complex goal-orientated tasks activate multiple regions within the PFC, thus complicating the ability to isolate the specific structures associated with different types of control functions (see Badre & D'Esposito 2009, for discussion). Further, hampering efforts to understand the functional and anatomical organisation is the relative rarity of individuals with circumscribed frontal lesions (Stuss et al., 1995; Stuss, 2011; Stuss & Alexander, 2007). Finally, efforts to delineate cognitive control are further made challenging by inadequate conceptualisations of control processes. Specifically, many current definitions of cognitive control are relatively broad, and lack concrete concepts that can be adequately operationalised (Stuss et al., 1995; Rabbit, 1997;

Tsuchida & Fellows, 2013). Consequently, many tasks held to measure control processes are multifactorial and contain a number of cognitive processes, a feature that limits interpretation and the isolation of specific functional and anatomical substrates (Stuss et al., 1995; Tsuchida & Fellows, 2013). To illustrate, consider the Symbol Digit Modalities Test (Smith, 1968, 1982), which requires an individual to substitute numbers with their corresponding geometric figures. Whilst this task serves as a prominent measure of divided attention, it also places heavy demands on a number of other cognitive processes, such as concentration, speed of information processing, memory, and visual tracking and scanning (Bate, Mathias, & Crawford, 2001; Smith, 1982). Other examples include the Tower of London Test, which assesses complex, higher-level planning (Shallice, 1982), and the Wisconsin Card Sorting Test (WCST), which evaluates set shifting and flexibility in the face of changing reinforcement contingencies. Each of these tasks arguably utilise a number of cognitive processes including attentional monitoring, sustained attention, motor-speed, spatial and working memory and inhibition. Accordingly, this sheer complexity may account for the low correlations often observed between tasks held to measure aspects of cognitive control performance and behaviour in naturalistic settings (Amieva, Phillips, & Della Sala, 2003; Bogod, Mateer & MacDonald, 2003; Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Chan, 2001; Chaytor, Schmitter-Edgecombe & Burr, 2006).

Taken together, the above limitations highlight the need to characterise cognitive control using more robust, theoretically-driven measures that involve few component processes. Doing so can enable for a more detailed evaluation of the functional and anatomical architecture of cognitive control, whilst also allowing for the adequate isolation of brain-behaviour relationships. The following sections will provide a brief overview of some of the more prominent theories of cognitive control.

Theoretical Accounts of Cognitive Control

Traditional accounts of cognitive control can be grouped into either unitary theories, or hierarchical accounts. Unitary theories emphasise the unified function of control processes, wherein cognitive control exists as a global network supported by a single entity in the PFC (Koechlin and Summerfield, 2007; Stuss et al., 1995; 2005). In contrast, hierarchical theories emphasise the different domains over which different types of cognitive control operate, which range from highly concrete (e.g., control of a specific movement) to increasingly abstract levels of control (Badre & D'Esposito, 2009). Accounts vary in the extent to which they emphasise different organisational aspects. Some emphasise the dimension of

concreteness (e.g. progressively more abstract operations are performed by progressively more anterior PFC structures); others emphasise the degree of domain specificity of the operations (e.g. less domain specific operations are performed by more anterior structures), whereas others place more emphasis on the nature of the operations being performed (e.g., maintenance versus manipulation of this information). Each of these accounts will be discussed in detail below.

General Purpose Unitary Accounts

One prominent theory based on a general purpose unified framework is Norman and Shallice's (1986) Supervisory Attentional System model. This model makes a primary distinction between routine and non-routine processes and behaviour (Stuss et al., 1995). Routine behaviours can be decomposed into three components: (i) *processing modules*; (ii) *schemata, (*representations of previously learnt and routine behaviours); and (iii) *contention scheduling*. In contrast, non-routine and complex behaviours are governed by the *Supervisory Attentional System* (*SAS*). In order for routine behaviours to occur, the appropriate schemata must first be activated, either by external inputs, such as environmental factors, or alternatively from internal inputs, such as perceptual representations or the output of recently engaged schemata (Norman & Shallice, 1986; Shallice, 1988; Stuss et al., 1995; 2005). Consider the example of spilt water on the floor; this would serve as an *external input*, which would initiate an internal representation, or schema, based on prior knowledge, regarding what is appropriate to do in that situation ("floor mopping"). The activation of this schema would then specify our *processing modules*, which would generate the behavioural response of mopping it up. Importantly, various alternative schemata may become activated at the same time; these schemata compete for selection via a process known as *contention scheduling*. In the case of very routine behaviours, one schema will receive greater activation than the others, and thus will "win" the competition of selection. However, non-routine, and complex operations require the intervention of a separate process known as the Supervisory Attentional System (SAS). The SAS comes into play in situations where there are no single routine schemas appropriate for our behaviour, when the most highly activated schema is inconsistent with our internal goals, and in situations that are novel, require error-correction, and necessitate planning or decision making (Norman & Shallice, 1986). Accordingly, the SAS operates to bias the activation of schemas in favour of the ones more consistent with the internally presented goal (a similar concept is Baddeley's central executive: see Baddeley, 1996; 1998; 2000, 2007).

From an anatomical perspective, the Supervisory Attentional System model - and other unitary accounts - suggest a core set of neural regions localised within the PFC operate as coherent unit, regardless of the specific information type (Millar & Cohen, 2001; Tsuchida & Fellows, 2013). Another model that takes a similar approach is that of Duncan (2010), which proposes the existence of a 'multiple demand network' that is engaged in the control of complex behaviour. This network is supported by regions in the bilateral ventrolateral prefrontal cortex, anterior insula, dorsolateral prefrontal cortex, the adjacent dorsal anterior cingulate cortex, mid-dorsolateral, mid-ventrolateral and the pre-supplementary motor area (Duncan, 2006; 2010; Tsuchida & Fellows, 2013). The model views the PFC as a "general computational resource" that is capable of functionally adapting in order to meet a range of cognitive demands and information types (Duncan & Miller, 2002, p. 289).

Another integrative theory is that described by Miller and Cohen (2001). According to this model, the primary role of the PFC is to actively maintain an individual's goals and the means to achieve them. These goal representations then operate in a domain-general manner to bias the flow of activation to other brain structures via the establishment of appropriate mappings between inputs, and internal and external states. Specifically, this model proposes that biasing signals operate across the neural regions to either enhance or inhibit the flow of activity along stimulus-response pathways depending on the whether the response is consistent or inconsistent with the current goal. As long as the goal representation remains sufficiently activated, it will be able to bias the flow of activation away from the inappropriate stimulus-response pathway and toward the appropriate one (Miller & Cohen, 2001). According to this theory, external cues activate internal representations within the PFC, which then determines the appropriate course of action. In unfamiliar circumstances, or those that elicit multiple courses of action, the PFC actively maintains the individual's goal and biases activation toward a specific stimulus-response pathway by establishing and integrating the appropriate mappings needed to perform the task (Miller & Cohen, 2001).

A central tennet of unitary frameworks is that disruption to these supervisory/ regulatory networks will result in a generalised pattern of deficits across all novel and complex behaviours (Andres & Van der Linden, 2001; Bouquet, Bonnaud, & Gil, 2003). Support for these unitary frameworks come from studies reporting that patients with damage to the PFC show consistent patterns of generalised, non-specific cognitive impairments on a wide range of complex tasks. Such generalised failures, often referred to as Dysexecutive Syndrome, include deficits on tasks involving inhibitory control, complex reasoning and

solving complex problems, detecting abstract concepts and rule changes, and the initiation, implementation, or planning upcoming responses (Andres & Van der Linden, 2001; Bouquet et al., 2003; Burgess & Shallice, 1997; Shallice, 1982). For example, Andres and Van der Linden (2001) compared the performance of 13 patients with focal frontal lesions and their age-matched controls on a number of tasks held to be sensitive to the frontal lobes, such as: the Tower of London test (Shallice, 1982), which requires participants to plan a sequence of moves in order to achieve a desired end goal; the Hayling test (Burgess & Shallice, 1996), which requires participants to complete sentences with an unexpected word (e.g. "the captain wanted to stay with the sinking… *banana*"); and the Brixton test (Burgess & Shallice, 1996), which examines participants' ability to extract rules from a series of visually presented symbols. Overall, the frontal patients exhibited slowed performance across both the Tower of London and Hayling tests, each of which offer very different measures of goal-orientated behaviour, with no significant differences according to the precise localisation of the PFC lesion (Andres &Van der Linden, 2001). Such findings appear to suggest that a common process may underpin the wide-ranging deficits observed following lesions to the PFC.

However, there a number of problems with this line of reasoning. Specifically, since many tasks used to examine "control" processes are highly complex and multi componential, it is perhaps not surprising that patients tend to fail on multiple tasks. Related to this, it is often difficult to determine the specific source of impairment given the multifactorial nature of many common tasks. Further, a number of studies conducted on both clinical and non-clinical populations, report low and often non-significant inter-correlations between performance on tasks held to measure cognitive functions (e.g. Burgess et al., 1998; Lehto, 1996; Levin et al., 1996; Lowe & Rabbitt, 1997; Robbins et al., 1998; & Schachar, Tannock, & Logan, 1993). For example, Godefroy and colleagues (1999) examined a group of individuals with frontal lesions, and observed impaired performance on the Wisconsin Card Sorting Test (a test of set-shifting and planning), yet intact performance on the Tower of Hanoi task, a task that is conceptually similar to the Tower of London task (Godefroy, Cabaret, Petit-Chenal, Pruvo & Rousseaux, 1999). In contrast, Andres &Van der Linden (2001) reported that whilst their sample of frontal patients showed impaired performance on the Tower of London and Hayling tests, they performed comparably with controls on the Brixton test. These observations provide evidence for the existence of differential patterns of cognitive impairment, and suggest that separable components may underpin control processes. Such findings call for a framework that is capable of capturing the integrated

activity of cognitive control, as well as its possible fractionation.

Hierarchical Accounts

An alternative position adopted in several more recent models is that cognitive control processes operate as a collection of separable but interdependent subcomponents, rather than a dedicated unitary system. These models propose that goal-orientated behaviour is organised hierarchically into distinct, yet inter-related functional and anatomical subregions and components (Badre & D'Esposito, 2007; 2009; Badre et al., 2009; Botvinick, 2008; Botvinick, Niv, & Barto, 2009; Koechlin, Ody, & Kouneiher, 2003; Koechlin & Summerfield, 2007; Sakai & Passingham, 2003). Most hierarchical models emphasise a dominance relationship, whereby increasingly anterior regions of the dorsolateral PFC are involved in increasingly abstract aspects of control (Badre & D' Esposito, 2009; Botvinick, 2007; 2008).

For example, one hierarchical framework, proposed by Koechlin & Summerfield, (2007), emphasises the dimension of time over which different types of control operate (Koechlin & Jubault, 2006; Koechlin et al., 2003). This "cascade model" proposes that behaviour is dependent on two types of information: information in the immediate environment, and conditional information that relies on both the immediate context and prior events. At the lowest-level, behaviour is determined by stimuli in the immediate environment, which elicits a relatively automatic response (for example, *answering your phone when it rings*). According to the model, this form of behaviour is controlled by the premotor cortex, but does not require the involvement of the prefrontal cortex.

The second lowest level of cognitive control, *contextual control*, requires that information from a stimulus must be integrated with other information from the current context to determine which behaviour should be performed (for example, *answering your phone if it rings only if you are in your own home*). According to Koechlin and colleagues (2007), this type of control involves posterior regions within the dorsolateral prefrontal cortex (Brodmann's area 8 and 44). The next level of cognitive control, called *episodic control*, governs situations where one's behaviour must be further modified by information from recent past events (for example*, your host telling you they are expecting an important phone call, so to answer their phone if it rings*). This information overrides the normal context-specific rule, and also requires the individual to retain the earlier instruction (this is the "episodic" element), whilst also invoking the instruction at the appropriate time. According to

the model, this type of episodic control is localised to the anterior dorsolateral regions of the prefrontal cortex (Brodmann's area 9 and 46), and encompasses more higher-level control processes. Finally, the highest level of control is referred to as *Branching control*, and involves switching flexibly between two different behaviours according to internal or external cues (for example, *your host telling you they are expecting an important call at 3pm, so you should answer their phone if it rings; however, once the call has taken place, you should no longer answer their phone*). This high-level control involves the coordination and engagement of multiple goals simultaneously by allowing one goal to be suspended and maintained in a temporary state (*answer the phone*), whilst another is currently being completed (*wait until the important phone call has finished*). Koechlin and Summerfield (2007) argue this level of control is localised to the anterior tip of the lateral prefrontal cortex (Brodmann's area 10) and is the basis for higher-level cognitive functions, such as task-switching and multi-tasking.

The cascade model makes important predictions regarding patient behaviour. It predicts that damage to any of the regions involved in the control hierarchy will affect only those behaviours that require control at the damaged level or higher. For example, damage to the posterior lateral prefrontal cortex would affect an individual's ability to modulate their behaviour according to the context (contextual control), but would not affect simple lower-level behaviours, such as those requiring sensorimotor behaviours. Moreover, damage to the anterior dorsolateral prefrontal cortex will not affect sensorimotor behaviours or contextual control, but will affect higher-level control, such as episodic and branching control (see Koechlin & Summerfield, 2007).

Another prominent hierarchical model of control places less emphasis on the time frame involved, and more on the degree of abstractness of the plan being created/implemented (Badre & D'Esposito, 2009; Botvinick, 2007; 2008). A central tenet of this model emphasises that any complex goal can be viewed as a hierarchy of sub-goals and actions. At the top of the hierarchy are broad abstract plans that could be implemented in various ways (e.g., make a salad). These are represented in the most anterior portions of the lateral PFC. This broader plan can then be broken down into more concrete plans (e.g., use cucumber and tomato in the salad), and then ultimately into a sequence of action plans supported by secondary motor cortex (e.g., hold cucumber in the left hand, and slice with the right hand). As we proceed from abstract to successively more concrete plans, there is a progressive shift from anterior to more posterior PFC regions. At each level in the hierarchy,

potential plans compete for selection via competition mechanisms; higher-level representations work to bias behaviour towards appropriate behaviours and away from inappropriate ones. Support for this theory comes from patient and neuroimaging studies that demonstrate greater involvement of more anterior regions as the rules governing behaviour on the task become more complex (Badre & D'Esposito, 2009; Badre, Hoffman, Cooney & D'Esposito, 2009; Christoff, Keramatian, Gordon, Smith & Madler, 2009).

Other prominent hierarchical theories emphasise the domain specificity of goal-orientated behaviour, rather than the degree of abstractness or the time frame involved. In particular, these accounts suggest that anterior prefrontal regions subserve domain general functions, such as monitoring, whilst more posterior frontal regions are involved in domain specific functions, such as inhibition of verbal responses specifically (for a recent review see Badre, 2008 and Badre & D'Esposito, 2009).

Regional Specialisation Models

Other theories place less emphasis on hierarchical aspect of control and more on the specific processing specialisations of various PFC subregions (Ridderinkhof, et al., 2004; Stuss et al., 1995; 2005; Tsuchida & Fellows, 2013). Some theories of this kind focus specifically on just one structure or type of process. For example, Botvinick and colleagues (2001) propose that medial prefrontal cortex, specifically the anterior cingulate cortex, may be specialised for monitoring conflict (Botvinick Braver, Barch, Carter, & Cohen, 2001; Ridderinkhof et al., 2004). This monitoring system anticipates situations where response conflict might be high, and subsequently recruits additional resources in order to re-adjust behaviour to minimise the chances of error. One of the behavioural manifestations of this readjustment process is post-error slowing: the observed increase in response times on a trial immediately following an error (Botvinick et al., 2001).

Stuss and colleagues (1995) provide a more comprehensive model of PFC function that emphasises the importance of regional specialisation. Drawing on the basic framework of Norman and Shallice (1986; Shallice, 1982), they propose that there are specific anatomical subregions of the PFC that subserve different broad components of behavioural and cognitive control, such as arousal, inhibition and monitoring (Stuss et al., 1995; 2005; see also Stuss, 2011). Specifically, these authors identify six different types of control processes: *energising schemata* (facilitation and allocation of arousal); *monitoring the level of activity in the schemata; inhibiting task-irrelevant schemata; adjusting contention scheduling; controlling*

('if this then that') logic; and *task-setting* (Stuss et al., 1995; 2005). It is further proposed that each of these processes are supported by distinct anatomical regions within the frontal lobes (e.g. Alexander, Stuss, Shallice, Picton, & Gillingham, 2005; Stuss et al., 2005). Below, we describe three of the most important of Stuss and colleagues' (1995; 2005) proposed component processes in detail: *energising schemata, inhibiting task-irrelevant schemata,* and *attentional monitoring.* These components were selected as they are arguably the most elemental processes, and thus provide an opportunity to isolate the most basic components of cognitive control. Consequently, insights from these three processes can provide a greater understanding of more complex component processes and cognitive operations.

Energising schemata. In Stuss and colleagues' (1995) framework, *energisation* is a broad term used to refer to the facilitation and allocation of arousal of the neural systems. It is proposed that schema, once activated, will gradually lose activation unless reactivated in some way (Stuss et al 1995; Stuss & Alexander, 2007). Such events may occur under conditions whereby complex, or repetitive stimuli occur at a slow or infrequent rate, leading to an increased susceptibility to competition from task-irrelevant stimuli (Stuss & Alexander, 2007). Under these conditions, the individual must not only *initiate* an appropriate response set, but must also *continuously activate* the relevant stimulus-response pairing to maintain that response set (MacPherson et al., 2010; Stuss et al 1995; 2005; Stuss, 2011).

Stuss and colleagues (1995; 2005) have operationalised their concept of energisation using a series of response time (RT) tasks that vary as to the number of choices to be made and the specific timing of events. For example, they compare and contrast performance on the following tasks: i) a Simple RT task (responding to a single stimulus with the same key each time); ii) a Choice RT task, where four different stimuli may appear (e.g., A, B, C and D), with one requiring a different response to others (e.g. press one key for 'A', and another for the remaining stimuli); and iii) a "Prepare" RT task, which is identical to the Choice RT task with the exception of a warning stimulus that is presented one or three seconds prior to the presentation of the target stimuli (Stuss et al., 2005). Importantly, across each of these tasks, stimulus presentation was manipulated with varying inter-stimulus intervals ranging from three to seven seconds. Specifically, Stuss and colleagues (2005) predicted that patients with impaired energisation will exhibit the following profile: slower overall response latencies, however latencies will become faster after a one second warning signal due to an increase in phasic alertness that serves to re-energise the schema. Overall, in a sample of 38 patients with focal frontal lesions, those with superior medial lesions were abnormally slow across the

trials, and were the only frontal group to show a significant difference of warning stimulus delay (slower responses in the 3s warning condition than in the 1s condition) - that is, these patients appeared to lose their energised response after a short delay.

Inhibiting task-irrelevant schemata. According to Stuss and colleagues (1995), the process of inhibition is called upon when automatic processes select schemata that are inconsistent with the current task goal, or where there are strong competing alternatives (Ridderinkhof et al., 2004; Stuss et al., 1995). This process has typically been operationalised through a Stroop paradigm (in which participants name the ink colour of the word whilst ignoring the word name itself). In this instance, inhibition is called upon when the word and ink colour do not match, and consequently, participants must utilise cognitive control to over-ride their automatic response of naming a word. Inhibition has also been operationalised through a modified Stroop paradigm. Indeed in an earlier study, Alexander, Stuss, Picton, Shallice, and Gillingham (2007) created a modified Stroop paradigm, wherein participants had to respond only when a specific conjunction of a stimulus and colour was presented. In this task, the target stimuli were a red X or a blue O, which appeared 25% of the time. The distractor stimuli included red O's and blue X's, as well as letters other than X or O. Overall, patients with lesions to the left dorsolateral region committed significantly more false positive errors on the distractor items, relative to the remaining frontal patients and healthy controls. Alexander and colleagues (2007) concluded that the left ventral-lateral region, which includes dorsolateral areas, is critical for establishing the contingent relationship of stimuli to responses, particularly under conditions that require continuous refreshing and suppression of more salient responses (see also Fassbender et al., 2004). Nonetheless, several lesion and neuroimaging studies suggest that there may be distinct modality-specific functions of inhibitory control, with non-verbal inhibition localised more strongly to right frontal regions (see Hamilton & Martin, 2005). For example, using event-related functional MRI (fMRI) in a sample of 14 neurologically healthy participants, a strong right lateralisation effect was observed when participants were required to withhold a prepotent motor response (see Garavan, Ross, & Stein, 1999; see also Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Floden & Stuss, 2006; Picton et al., 2006). Such findings led Garavan and colleagues (1999) to conclude that response inhibition is lateralised to right frontal regions of the PFC.

Attentional Monitoring. Attentional monitoring is a relatively broad concept encompassing a variety of processes that act to monitor the state of the internal and external environment against the current goals of the task (Botvinick et al., 2001; Ridderinkhof et al.,

2004; Stuss et al., 1995; 2005; Stuss, 2011). This may include monitoring the environment for responses or material that may be relevant to current goals; monitoring the level and timing of ongoing activity in task-relevant schema; anticipating potentially important upcoming stimuli, and how to respond to them; and monitoring one's own recent behaviour and performance effectiveness – for example, detecting errors when there is a discrepancy between the desired outcome and our performance (Ridderinkhof et al., 2004; Stuss et al., 2005; Stuss, 2011). Accordingly, it is these operations that allow for flexible adjustments of behaviour to ensure goal-appropriate behaviour (Ridderinkhof et al., 2004).

Stuss and colleagues (1995; 2005) operationalise this concept in various ways. They suggest that one indicator of effective monitoring is the foreperiod effect: the finding that in a RT task with variable inter-stimulus intervals (ISI's), participants should exhibit faster response latencies following the longer ISIs, compared to when the task involves the same ISI, which is consistent throughout the task. According to Stuss and colleagues, responding rapidly to stimuli following long ISI's requires effective *energisation* (see above). In addition, being able to modulate response times when ISIs are *variable* indicates the participant is successfully utilising the ISI to anticipate and prepare for the next stimulus, and finally, is taking into account the broader characteristics of the trial (the fact that sometimes, there may be little time until the next stimulus so they need to act quickly). Conversely, failures in this monitoring system can give rise to a reverse foreperiod effect, as evidenced by slower response latencies following longer, relative to shorter intervals.

Stuss and colleagues (2005) reported that individuals with lesions to right lateral PFC were most likely to exhibit these reversed foreperiod effects. In contrast, these individuals demonstrated intact performance on a fixed ISI task, suggesting the failure on the foreperiod task was indeed due to poor monitoring specifically (see Alexander et al., 2005).

Moreover, this group of researchers also propose that poor monitoring will be associated with particular types of error patterns on verbal fluency tasks – specifically, high rates of perseverative errors (repetitions of the same response within the same trial block). In this context, perseverative errors can only be effectively prevented if the individual is keeping careful track of their recent responses and monitoring their self performance (Reverberi, Lavaroni, Gigli, Skrap, & Shallice, 2005; Shao, Janse, Visser, & Meyer, 2014). Again, this pattern of performance has been associated with lesions to right frontal lateral regions. Patients in this group have been found to produce twice as many perseverative errors on

verbal fluency tasks, relative to those with left-hemispheric lesions (Stuss et al., Alexander, Palumbo, Buckle, Sayer, Pogue, 1994).

Chapter 2: Long-Term Language Outcomes in Tumour Patients
Introduction

According to recent reports, the worldwide incidence of primary neurological tumours is 3.4 cases per 100,000 people, with over 26,000 primary malignant and 53,000 primary non-malignant tumours expected to be diagnosed in the United States for the year 2017 (Ostrom et al., 2015, 2016, 2017). Recent estimates suggest nearly 700,00 people in the United States are living with a primary brain tumour, with approximately 17,000 losing their lives annually (Ostrom et al., 2017). Within the New Zealand context, estimates suggest there are between 100-120 new cases of malignant brain tumours annually (Ministry of Health, 2010); indeed, this incidence was projected to rise for the year 2016, with a 24 percent increase for males and 8 percent for females (Ministry of Health, 2010). Given these projected increases, it is necessary to evaluate the long-term functional impact and clinical outcomes associated with this aetiology. Indeed, one crucial function that is particularly vulnerable to impairment in cases of neurological tumours is language. Given the importance of language for everyday social interaction, there is an especially strong need to better understand how it can be impacted by brain tumour and its treatments (Bartha, Knosp, Pfisterer, & Benke, 2000; Finch & Copland, 2014; see also Davie et al., 2009; Faulkner, Wilshire, Parker, & Cunningham, 2017; Sanai, Mirzadeh, & Berger, 2008; Thomas, O'Conner, & Ashley, 1995; Whittle, Pringle, & Taylor, 1998).

Brain Tumours: An Overview. A neurological tumour consists of a solid mass of tissue within the brain, which is formed through abnormal and uncontrolled cell division. Primary neurological tumours are those that develop within tissue located within the cranium, such as: the neuroepithelial tissue associated with the brain itself (e.g. astrocytomas, oligodendrogliomas, ependymomas, glioblastomas), the cranial nerves (e.g. schwannoma), blood vessels (e.g. hemangioblastoma), pineal gland (e.g. pineocytoma), skull (e.g. pituitary adenoma), and the connecting tissue surrounding the brain (e.g. meninges: meningioma) (Ostrom et al., 2014). Brain tumours can also occur because of metastases of malignant cells from other organs (e.g. a primary tumour originating elsewhere in the body that metastasises to become a secondary tumour in the brain).

According to recent estimates, the most common form of primary neurological tumour is glioma (50.4%), followed by primary meningioma (20.8% - 36.6%), and pituitary adenomas (15%) (Park, Kim, Sade & Lee, 2009; Ostrom et al., 2017). Malignancy classifications of neurological tumours are categorised according to the World Health

Organisation (WHO) grading system, which based on the histological features of a tumour, differentiates between low malignancy grade (e.g. primary meningioma and low-grade glioma) and high malignancy grade (e.g. glioblastoma and high-grade glioma) (Bosman, Carnerio, Hruban & Theise, 2010). This grading system dictates that low-grade tumours (grade I and II) tend to be slow growing and have a low proliferative potential, with a high possibility of cure following surgical resection. Conversely, high-grade tumours (grade III and IV) are generally mitotically active and associated with advanced and rapid disease progression, with treatments often encompassing a combination of surgery and chemotherapy and/or radiation therapy (Louis., et al 2007). The specific histologic features used for each grade are presented in Table 2.1. Accordingly, it is these sources of tumour diversity (e.g. low vs high proliferative potential and slow vs fast-growing) that typically dictate the course, progression, and treatment options in neurological tumour patients. As will be discussed in subsequent sections, unique pathological characteristics distinguish neurological tumours from other aetiologies; this includes their histology, growth profiles, distinct localisation patterns, and potential for compression and/or displacement of surrounding tissue.

Table 2.1.

The WHO grading of Central Nervous System tumours (see Louis et al., 2016).

WHO grade I	Tumours with low proliferative potential, typically a slow-growing and discrete nature, and the possibility of cure following surgical resection without chemotherapy or radiation (e.g. meningioma)
WHO grade II	Tumours that are generally infiltrating and low in mitotic activity but may reoccur at a later time (e.g. diffuse astrocytoma)
WHO grade III	Tumours with histologic evidence of anaplasia and malignancy, generally in the form of mitotic activity, clearly expressed infiltrative capabilities. Often advanced and rapid disease progression (e.g. anaplastic astrocytoma)
WHO grade IV	Tumours that are mitotically active, necrosis-prone, and often associated with rapid preoperative and postoperative disease progression (e.g. glioblastoma)

Brain Tumours and Language Assessment

Conventional language assessment batteries have largely been developed from the framework of the classical model of language (Broca, 1861; Lichtheim, 1885; Wernicke, 1874). This model draws heavily on evidence from patients with post-stroke aphasia. Specifically, the classical model proposes a very simple neuroanatomical framework of linguistic functions, whereby two core regions are emphasised in the production and comprehension of language: a region within the left inferior frontal gyrus, referred to as Broca's area, and a region within the left superior temporal gyrus referred to as Wernicke's area (see Figure 2.1) (Broca, 1861; Lichtheim, 1885; Wernicke, 1874). According to this framework, Broca's area is necessary for the production of speech and more specifically, the storage of articulatory representations of words, whilst Wernicke's area is necessary for the receptive aspects of speech, and is considered the storage site for auditory forms of words (Broca, 1861; Damasio, 1998; Wernicke, 1874). In addition, a disruption to the white fibre tracts that connect these classical areas has been associated with a specific form of aphasia that affects the capacity to repeat words and phrases, known as Conduction aphasia (Geschwind, 1965; Lichtheim, 1885; Wernicke, 1874). Finally, a number of secondary types of Transcortical Motor Aphasia, and Transcortical Sensory Aphasia are identified, each of which are associated with distinct patterns of linguistic impairment.
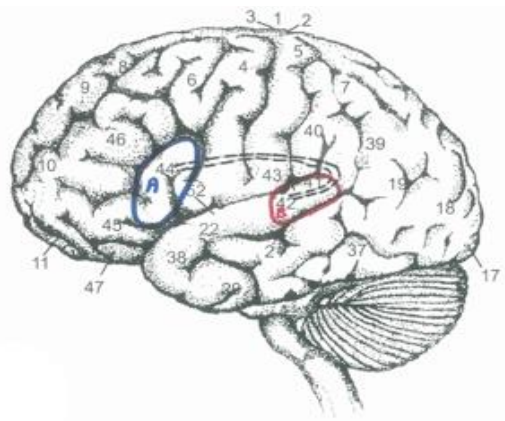
Figure 2.1. A representation of the classical model of language, which dictates two prominent regions critical for language. Broca's area is represented by the blue circle (A), whilst Wernicke's area is represented by the red circle (B). The arcuate fasciculus connects the two regions via a bundle of white fibre tracts. This diagram is courtesy of Wilshire (2014).

Current language assessment batteries, such as the Western Aphasia Battery (WAB;

Kertesz, 1982), Boston Diagnostic Aphasia Examination (BDAE Goodglass, Kaplan, & Barresi, 2001), and Aachen Aphasia Test (AAT; Huber, Poeck, Weniger, & Willmes, 1984) extend upon this basic framework and the core principles of the classical model of language. Such assessments classify individuals based on the quality of their spontaneous speech, and also their performance on tasks involving comprehension, picture naming and repetition. These assessment batteries provide operationalised definitions of the classical aphasia syndromes of Broca's, Wernicke's, Conduction and Transcortical aphasias, but also group them into larger and broader categories of *fluent* and *non-fluent* aphasia. Specifically, *fluent aphasia* includes those syndromes that are characterised by fluent language production, but compromised comprehension abilities (e.g., Wernicke's aphasia, Conduction aphasia, Transcortical sensory aphasia), whilst *non-fluent aphasia* includes those syndromes characterised by fragmented, sparse and/or effortful language production, but generally intact comprehension (e.g., Broca's aphasia, Transcortical motor aphasia). In addition to these classical syndromes, the above language assessment batteries also include more contemporary aphasia categories, such as *Global* aphasia, considered a severe form of non-fluent aphasia characterised by both production and comprehension deficits, and *Anomic* aphasia, a relatively mild form of fluent aphasia, predominately characterised by selective difficulties retrieving and expressing words.

Although these types of assessment batteries have been developed for, and validated primarily on stroke populations, they are often utilised in other neurological populations, including tumour patients (Poeppel & Hickok, 2004; Whittle et al., 1998; Wu et al., 2011). Indeed, Wu and colleagues (2011) recently reported that both the WAB (Kertesz, 1982) and BDAE (Goodglass et al., 2001) were amongst those commonly used to investigate language functioning in neurological tumour patients during the acute surgical phase (see also Davie et al., 2009). These studies have found that the language profiles of tumour patients differ from those of stroke patients in several major ways. For example, Whittle and colleagues (1998) assessed an undifferentiated sample of 40 patients undergoing surgery for a left-hemispheric tumour. Patients were assessed both pre- and postoperatively. Overall, 63% of patients scored below the normal range on the Aphasia Quotient on the WAB, whilst a high proportion of patients also scored below the normal range on the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983). Similar results were obtained by Wacker and colleagues (2002). These researchers assessed a sample of 100 individuals with primary or metastatic neurological tumours using the Aachen Aphasia Test (AAT; Huber et al., 1984), which

classifies language according to the primary aphasic subtypes (Wernicke's, Broca's, Global, and Amnestic), and was developed and standardised primarily for post-stroke aphasia. The sample included both left and right hemisphere cases. During the pre-operative assessment, at least 50% of those with left hemisphere tumours, and 36% of those with right hemisphere tumours, were classified as impaired on the AAT, which is defined as a deficit on at least one of the five subtests (Wacker, Holder, Will, Winkler & Ilmberger, 2002).

Despite the high rates of impairment on conventional language assessments, only a small number of studies have directly compared the language profiles between stroke and tumour populations (e.g. Anderson, Damasio, & Tranel, 1990). Language profiles in post-stroke aphasia are commonly characterised by moderate to severe linguistic impairments, whilst patients with neurological tumours typically exhibit relatively selective language deficits, which can be difficult to detect in the context of everyday conversation (e.g. Anderson et al., 1990; Haas, Vogt, Schiemann, & Patzold, 1982; Haglund, Berger, Shamseldin, Lettich, & Ojemann, 1994; Miceli, Caltagirone, Gainotti, Masullo, & Silveri, 1981; Recht, McCarthy, O'Donnell, Cohen, & Drachman, 1989; Whittle et al., 1998). For example, in a recent study, Davie and colleagues (2009) assessed a sample of 63 patients with left-hemisphere malignant tumours using the WAB (Kertesz, 1982). Each patient had undergone surgical resection for malignant tumours within 35 days of testing administration. Overall, patients exhibited low rates of global aphasia (3%), and high rates of anomic aphasia (49%), suggesting relatively selective impairments. This profile contrasts markedly with that of post-stroke aphasia, which typically presents with higher rates of global aphasia (20-30%), yet comparatively low rates of anomic aphasia (9-28%) (see Kauhanen et al., 2000; Pashek & Holland, 1988; & Pedersen, Vinter, & Olsen, 2004). These differential rates are consistent with that reported in an earlier study by Anderson and colleagues (1990), who evaluated language functioning in a sample of 17 undifferentiated neurological tumour patients and 17 stroke patients. Patients were assessed on both the Multilingual Aphasia Battery (which classifies impairment according to traditional aphasia categories; Benton, 1969) and the Boston Diagnostic Aphasia Examination Reading Sentences and Paragraphs subtest (Goodglass & Kaplan, 1983). Despite tumours being matched with vascular lesions for size and location, the two aetiologies exhibited considerable differences in their linguistic profiles. Specifically, left hemisphere stroke patients displayed greater impairment than the tumour patients on the BDAE Reading subtest, and on all subtests of the Multilingual Aphasia Examination (Anderson et al., 1990). Similar findings have been described by Miceli and

colleagues (1981), who reported that individuals with post-stroke aphasia exhibited proportionately greater rates of non-fluent aphasia than those with brain tumours, whilst those in the latter group demonstrated greater rates of anomic aphasia. Taken together, these findings suggest that differential linguistic profiles exist between post-stroke and neurological tumour patients, with stroke associated with greater rates of more generalised and severe language impairments relative to neurological tumours. Whilst the mechanisms underpinning these differences remains unclear, the greater rates of anomia in tumour populations may indeed reflect a mild form of aphasia that fails to contain enough features to qualify as a major aphasic category, and thus tends to default to anomia. Below, we briefly discuss some factors that may account for the differential language profiles.

Given the very different mechanisms underlying stroke and tumour, it is perhaps not surprising that patients' language profiles are distinctly different. Specifically, the neural regions most affected by vascular lesions are those that lie within the occluded or ruptured artery, with specific regions more vulnerable to vascular lesions than others (Miceli, Capasso, Monti, Santini, & Talacchi, 2012). Indeed, an occlusion or rupture to the middle cerebral artery (MCA) will almost always result in damage to the left inferior frontal gyrus (an area implicated in Broca's aphasia), whilst lesions to the posterior branches are commonly associated with damage to the left superior temporal gyrus (an area implicated in Wernicke's aphasia; Miceli et al., 2012); it is these site-specific patterns that are likely to account for lesion location serving as a primary determinant of language dysfunction in stroke populations (Taphoorn & Klein, 2004). Neurological tumours on the other hand tend to have greater localisation disparity, particularly when involving the interstitial tissue (for discussion, see Miceli et al., 2012). Importantly however, the clinical profile exhibited by tumour patients varies not just according to the location and size of the lesion, but also by tumour type and grade. Indeed, a number of studies have reported it is the grade of tumour opposed to the location that serves as the primary predictor of language difficulties and outcomes in tumour populations (Bello et al., 2007; Ilmberger, Ruge, Kreth, Briegel, Reulen, & Tonn, 2008; see discussion below). Importantly though, the role of tumour grade and type has not been adequately investigated in the literature (Noll, Sullaway, Ziu, Weinberg, & Wefal, 2014).

Distinct time courses may also play a significant role in the differential language profiles that exist between tumour and stroke patients. Indeed, vascular lesions, such as those resulting from stroke, result in tissue necrosis – a phenomenon that refers to the sudden

disruption to the grey and white matter tracts that form pathways throughout the brain and allow separate regions to communicate (Davie et al., 2009; Miceli et al., 2012). Conversely, the growth rate of neoplastic lesions is often much slower, thus allowing for the gradual displacement of surrounding healthy cortical tissue (Duffau et al., 2003); this displacement ultimately provides an opportunity for functional reorganisation, and the subsequent development of compensatory strategies to occur (Duffau et al., 2003; Duffau, 2007; Finch & Copland, 2014; Miceli et al., 2012). It is these strategies that likely accounts for the finding that tumour growth can often be quite advanced before functional impairments become evident (Anderson et al., 1990; Davie et al., 2009; Desmurget, Bonnetblanc, & Duffau, 2007; Miceli et al., 2012). Accordingly, a detailed and sufficiently sensitive assessment is needed to detect the subtle functional impairments that are often present in neurological tumour populations.

Different patterns of recovery are also evident between stroke and tumour populations. Specifically, a number of studies report an initial decline in tumour patients' language functions immediately following surgery, however considerable recovery is generally evident after three months (Finch & Copland, 2014; Sanai et al., 2008; Wu et al., 2011). Conversely, in post-stroke aphasia, some degree of spontaneous recovery is typically evident within 8-12 week's post-onset, however such improvements generally plateau within one year, with few functional improvements evident thereafter (Berthier, 2005; Shafi & Carozza, 2012).

Given that distinct mechanisms exist between post-stroke and neurological tumour patients, conventional aphasia assessments, validated largely on post-stroke populations, may not be optimal for assessing language impairment in those with neurological tumours. This was demonstrated in a study by Påhlson, Ek, Ahlstrom and Smits (2003), who compared language and cognitive functions in 24 patients with low-grade glioma using three distinct types of assessment approaches. The *neurologist's assessment* included a standard aphasia assessment - the Boston Aphasia Severity Rating Scale, adapted from the BDAE - and several other measures, including the Edinburgh Functional Impairment Test, and the Williams Delayed Recall Test. The *neuropsychologist's* assessment included a range of tests of attention, reasoning, and visual-spatial organisation and judgement. Finally, the *self-report assessment* included a range of questionnaires regarding patients' own perceptions of their general level of disability and cognitive functioning. Overall, significant differences were evident between the three approaches, with the comprehensive neuropsychological

assessment detecting moderate to severe cognitive impairments, which were not evident on either the self-report measures or the neurological evaluations. Upon further inspection, the most common impairment detected by the neuropsychologist's assessment was verbal memory, whilst language was one of the least detected by the neurologist's assessment (Påhlson et al., 2003). These results suggest that conventional assessment tools validated for stroke populations may not adequately capture the mild language impairments, nor the involvement of higher-level linguistic functions in neurological tumour populations (Davie et al., 2009). Indeed, this view has been endorsed by a number of previous researchers (De Witte & Mariën, 2013; Meyers & Brown, 2006; Miceli et al. 2012; Talacchi, Santini, Savazzi & Gerosa, 2011).

Recently, comprehensive neuropsychological assessment protocols, tailored specifically for tumour patients, have started to emerge. Indeed, the recent Milano-Biocca Battery was developed to assess the following three domains in neurological tumour populations: language, memory, and executive functions. To investigate language, the following tasks were administered: *category* and *letter fluency, action* and *object naming, naming famous people, picture-word matching, naming by description,* and *real word* and *non-word sentence repetition* (Papagno, Casarotti, Comi, Gallucci, Riva, & Bello, 2012). Overall, 226 patients with neurological tumours were assessed at three time-points: pre and post-operatively, and at three months' post-surgery. The following tasks were most sensitive at detecting linguistic deficits, with preoperative impairments ranging between 60-79 percent: *category* and *letter fluency, action* and *object naming,* and *naming famous people*. Further, after categorising patients according to left and right frontal and temporal localisations, tumour location was found to be a strong predictor of language functioning, with left temporal tumour patients exhibiting greater impairments on category fluency and naming tasks (e.g. objects and famous people), whilst those with left frontal tumours exhibited more impairment on a letter fluency task (Papagno et al., 2012).

Another recent language protocol tailored specifically for tumour populations is the Dutch Linguistic Intraoperative Protocol (DuLIP; De Witte, Satoer, Robert, Colle, Verheyen, Visch-Brink, & Mariën, 2015). The DuLIP is based on a linguistic framework that assesses the following broad domains of linguistic processes: *semantic* (word and picture matching, a word association task, two category fluency tasks, and various sentence meaning judgement tasks), *syntactic* (verb generation, word production tasks, and a syntactic judgement task), *phonological* (word repetition, letter fluency and various phonological discrimination tasks),

and *articulation*. Preliminary data is available for five patients with tumours in eloquent language areas, across three stages: preoperatively (six months pre surgery), intra-operatively (during electrical stimulation), and postoperatively (six months' post-surgery) Overall, verb generation, action naming and fluency were considered to be the most sensitive at detecting impairments in this sample (De Witte et al., 2015).

Both the DuLIP and the Milano-Biocca Battery provide an arguably more sensitive and comprehensive assessment of language functions that are tailored specifically for neurological tumour patients. However, one limitation is that many of the tasks in these assessments are highly complex, and are therefore likely to engage a wide range of verbal and non-verbal cognitive skills. To demonstrate this, consider the verbal fluency task. In this task, participants are asked to generate as many words that belong to a particular semantic category (e.g. fruit), or that start with a given letter (e.g. "F") in 60 seconds. Successful completion of this task not only requires a number of verbal processes, but also requires participants to formulate a strategy for searching through their memory for appropriate exemplars, internally-generate the most appropriate verbal responses, sustain their attention, suppress any irrelevant or competing schema and intrusions, and monitor responses to ensure inappropriate or perseverative errors are not produced (Alexander, 2006; Shao, Janse, Visser, & Meyer, 2014). Recently, researchers from our group developed a new protocol for assessing language in tumour patients, called the Brief Language Assessment for Surgical Tumours (BLAST; Faulkner, 2015; Faulkner et al., 2017). The BLAST draw on current cognitive and neuropsychological theory to identify the core cognitive operations deemed necessary for effective single word production. This approach offers a systematic and theory-driven method for assessing language competence in the minimum time possible. To ensure brief administration time, the BLAST includes several tasks that evaluate multiple linguistic skills, and manipulates different properties of the stimulus words in order to tease the different skills and processes apart (see Table 2.2). Importantly, one advantage of the BLAST is its short administration time of between 25 and 40 minutes, depending on the patient's overall level of functioning.

The Brief Language Assessment for Surgical Tumours.

As noted above, most conventional aphasia assessments are based on a very coarse-grained model of language function that identifies only broad components, such as comprehension, repetition, and production/fluency. However, as discussed in Chapter 1, recent advances in cognitive theory offer a finer-grained theoretical framework of language

functions, which may prove more effective at assessing neurological tumour patients. Drawing on such theory is the Brief Language Assessment for Surgical Tumours (BLAST; Faulkner, 2015; 2017). Specifically, the BLAST is designed to measure eight core cognitive processes described previously: Auditory Word Recognition, Accessing Semantic Knowledge, Lexical Selection, Phonological Encoding, Verbal Short-term Memory, Goal-Driven Response Selection, Verb Retrieval, and Articulatory Motor Planning. The assessment comprises eight different tasks, which are summarised in Table 2.2. As can be seen from the table, some tasks are designed in such a way that more than one measure can be extracted from them, enabling us to measure more than one cognitive skill within a single task. In this way, administration time for the test is kept to a minimum.

Table 2.2 provides a summary of each of the eight cognitive processes assessed in the BLAST, the tasks used to derive each cognitive skill, the cognitive and neuroanatomical profile that has been associated with an impairment to each, and finally, the specific features of that profile that were assessed in the BLAST.

Table 2.2.

Individual summaries and a description of the language profile associated with a selective impairment for each core cognitive skill (see Faulkner et al., 2017)

Cognitive Skill	Tasks used to derive each cognitive skill	Population most commonly associated with selective impairment	Language profile associated with hypothesised selective impairment	Specific aspects of profile assessed in the BLAST	Neural substrate as identified in large group lesion analyses
Auditory Word Recognition	Picture-word verification Real-word repetition	Fluent aphasia associated with damage to the superior posterior temporal regions (Buchman, Garron, Trost-Cardamone, Wichter & Schwartz, 1986; Poeppel, 2001)	Disproportionately impaired auditory word comprehension. Patients exhibit difficulty with word-to-picture matching tasks on phonologically-related items. Difficulty discriminating between auditory lexical items. Deficits on single word repetition and reverse length effect as evidenced by poorer performance on shorter, relative to longer words (e.g., Howard & Franklin, 1988; Wilshire & Fisher, 2004).	Deficits on single-word repetition. Confusions between phonologically-related items on picture-word verification. Reverse length effect on real-word repetition	*Left posterior temporal-parietal* damage (esp. *posterior superior temporal gyrus)* associated with inability to discriminate non-word minimal pairs in 10 patients with Wernicke's aphasia (Robson, Sage & Lambon Ralph, 2012)
Accessing Semantic Knowledge	Picture Naming Picture-word Verification Category and Letter Fluency	Semantic dementia associated with damage to degeneration of the left anterior temporal lobe (Hodges et al., 1999; Jefferies & Lambon Ralph, 2006)	Semantic errors on picture naming and matching, particularly on within-category substitutions (e.g. *deer ->* camel). Greater errors on low frequency, relative to high frequency, items. Greater errors on word-to-picture matching tasks when distractors are semantically related. Disproportionately poor performance on category relative to letter fluency.	Semantic errors on picture naming. Disproportionately poorer performance on category, relative to letter fluency. Greater confusion and errors between semantically related items on picture-word verification.	*Left anterior temporal lobe* damage associated with semantic errors in picture naming in a large group of patients with left hemisphere lesions (Schwartz et al., 2009).

Cognitive Skill	Tasks used to derive each cognitive skill	Population most commonly associated with selective impairment	Language profile associated with hypothesised selective impairment	Specific aspects of profile assessed in the BLAST	Neural substrate as identified in large group lesion analyses
Lexical Selection	Picture Naming Picture-word verification Letter and category fluency	Anomic aphasia associated with left posterior temporal lesions (Antonucci, Beeson, Labiner & Rapcsak, 2008; Foundas, Daniels & Vasterling, 1998; Lambon Ralph, Sage & Roberts, 2000; Raymer et al., 1997).	Errors on picture naming, including omissions and semantic errors, particularly semantic associates (e.g. *needle* -> "thread"). Word frequency effect on picture naming. Word-to-picture matching better than naming, although may not be normal (e.g., Foundas et al., 1998; Raymer et al., 1997). Word repetition superior to naming.	Frequency effect in picture naming. Disproportionately high production of omission and semantic errors. Relatively well-preserved performance on picture-word matching tasks.	Hyperfusion in *posterior inferior and middle temporal cortex* and *posterior angular gyrus* is most strongly correlated with impaired naming, but preserved picture-word verification in acute stroke patients (Deleon et al., 2007).
Phonological Encoding	Picture Naming Real-word and non-word repetition Articulatory agility task	Conduction aphasia associated with damage to posterior temporal and/or inferior parietal lobe (Bartha & Benke, 2003; Wilshire, 2002)	Phonemic paraphasia in picture naming, real-word reading and real-word repetition. Performance on picture naming is influenced by word length as well as frequency. Spontaneous speech is fluent, but contains phonemic paraphasias on longer and/or less common words. No effect of syllabic complexity on production accuracy (see esp., Romani, & Galluzzi, 2005). Auditory comprehension relatively preserved.	Strong length effect on picture naming and disproportionately high production of phonological errors. Poor performance in single non-word repetition. Relatively well-preserved articulatory agility.	*Left superior temporal and inferior parietal* damage associated with phonemic paraphasias on picture naming in a large left hemisphere lesioned sample (Schwartz, Faseyitan, Kim, & Coslett, 2012). Damage to *postcentral gyrus, inferior precentral gyrus, andsupramarginal gyrus* associated with impaired word repetition in a large left hemisphere lesion population (Baldo, Katseff & Dronkers, 2012).

Cognitive Skill	Tasks used to derive each cognitive skill	Population most commonly associated with selective impairment	Language profile associated with hypothesised selective impairment	Specific aspects of profile assessed in the BLAST	Neural substrate as identified in large group lesion analyses
Verbal Short-Term Memory	Non-word repetition	Conduction aphasia associated with damage to posterior and temporal and/or inferior parietal lobe (Baldo, Klostermann & Dronkers, 2008; Vallar & Shallice, 1990).	Impaired performance on verbal span tasks. Absence of recency effect on span tasks. Non-word repetition impaired, but real-word repetition may be normal. Poor sentence repetition; responses may paraphrase the original sentence (see esp., Baldo et al., 2008). Impaired comprehension of complex sentences (Baldo et al., 2008; Papagno, Cecchetto, Reati & Bello, 2007; Pettigrew & Hillis, 2014).	Poor performance on non-word repetition.	Damage to a region extending *from left superior and middle temporal gyrus to inferior parietal lobe* associated with (i) poor performance on non-word repetition; and (ii) poor performance on span tasks, in a group of left hemisphere stroke patients (Baldo et al., 2012).
Goal-Driven Response Selection	Stroop task

Verb generation

Letter fluency

Picture naming | Non-fluent aphasia associated with lesions to the left inferior frontal gyrus (Jefferies & Lambon Ralph, 2006; Martin & Freedman, 2001; Schnur, Schwartz, Brecher & Hodgson, 2006; Thompson- Schill et al., 1998). | On verb generation tasks, disproportionately impaired on items with multiple response possibilities (e.g., *rope* -> tie, cut, hang, pull). Performance on naming tasks declines when semantically related pictures are repeatedly presented (Schnur et al., 2006).

Impaired letter fluency (Robinson, Shallice, Bozzali & Cipolotti, 2010; Speer & Wilshire, 2014). Disproportionate Stroop effect on incongruent items (Hamilton & Martin, 2005; Scott & Wilshire, 2010). Difficulty producing utterances containing meaning-related words (Freedman, Martin & Biegler, 2004; Speer & Wilshire, 2014). | Impaired accuracy performance on tasks that require the resolution of conflict (Stroop), selection from amongst multiple alternatives (verb generation), or strategically driven lexical search (letter fluency). | Damage to a region extending *from left inferior frontal gyrus to the inferior parietal lobe* associated with poorer letter than category fluency in a left hemisphere stroke sample (Baldo, Schwartz, Wilkins & Dronkers, 2006).

Damage *to left inferior frontal gyrus* predicted the size of the Stroop effect (in response latencies) in patients with frontal lobe lesions (Tsuchida & Fellows, 2013). |

Cognitive Skill	Tasks used to derive each cognitive skill	Population most commonly associated with selective impairment	Language profile associated with hypothesised selective impairment	Specific aspects of profile assessed in the BLAST	Neural substrate as identified in large group lesion analyses
Verb Retrieval	Verb generation Picture naming	Most commonly non-fluent aphasia associated with damage to anterior language regions (although in some cases have more posterior lesions: Mätzig, Druks, Masterson & Vigliocco, 2009).	Disproportionately poorer performance on action naming than on object naming tasks. Particularly with anterior lesions, accuracy on sentence production declines with the number of obligatory verb arguments (see esp., Bastiaanse & Jonkers, 1998; Collina, Marangolo & Tabossi, 2001; Luzzatti et al., 2002; Thompson, 2003).	Disproportionately poorer performance on high-strength condition of the verb generation task compared to naming pictures of high frequency.	Damage to the left inferior frontal gyrus associated with disproportionately poor performance on an action naming task in a sample of 20 left hemisphere stroke patients (Piras & Marangolo, 2010).
Articulatory Motor Planning	Articulatory Agility task	Apraxia of speech associated with damage to left insula and/or surrounding regions (Josephs et al., 2006; Ogar et al., 2007; Wertz, LaPointe & Rosenbek, 1984).	Significantly reduced speech rate. Articulatory inconsistency on repeated productions of the same utterance, particularly multisyllabic word, alternating sequences (e.g. *pataka*) and sentences (see Ogar et al., 2007). Disproportionate difficulty on complex syllables compared to simple ones (see esp., Romani, & Galluzzi, 2005).	Poor latency performance on BDAE articulatory agility task (involves repeated productions of items that vary in syllable length (e.g., *thanks – huckleberry*).	Damage to *anterior insula, inferior frontal gyrus and anterior temporal lobe* associated with low scores on articulation subtests of the Aachen Aphasia Test in a large group of left hemisphere stroke patients (Henseler, Regenbrecht & Obrig, 2014).

In a recent study, Faulkner and colleagues (2017) administered the BLAST to a sample of 49 undifferentiated tumour patients during the acute preoperative phase (one day prior to surgery) (see Chapter 2 for the full administration procedures). Overall, the core skills approach revealed considerable sensitivity, with 65% of patients scoring below healthy controls on at least one core skill. To assess anatomical specificity of each of the core skills, patients were then categorised into one of the following six groups, based on their lesion location: *left frontal* (16), *left parietal* (9), *left temporal* (7), *right frontal* (12), *right parietal* (1), and *right temporal* (4). The core skills approach revealed finer-grained anatomical specificity, with left temporal patients performing significantly more poorly on lexical selection and accessing semantic knowledge, relative to the remaining patients. Further, patients with left temporal and/or left parietal lesions also exhibited significantly poorer performance on phonological encoding relative to the remaining patients. Finally, those with left frontal tumours performed significantly worse on articulatory motor planning relative to the other groups. Such observations were consistent with current cognitively-motivated theories of language (see Chapter 1), and suggests that decomposing linguistic functions according to a core skills approach allows for specific conclusions to be drawn regarding finer-grained linguistic processes, and their corresponding anatomical correlates.

Purpose of the Current Investigation

The BLAST is proving to have good capability at both capturing the variation amongst tumour patients and discriminating between different types of cognitive profiles. However, we still know very little about the relationship between patients' preoperative performance on the BLAST and how they perform following surgery. It is of value to examine patients' performance not just immediately after surgery, but also several months after surgery. This is because the immediate post-surgical phase is likely to induce complications including swelling, neural displacement, fatigue, medication effects, and psychological factors including depression and anxiety (Finch & Copland, 2014; Heimans & Reijneveld, 2012; Papaioannou, Fraidakis, Michaloudis, Balalis, & Askitopoulou, 2005; Talacchi et al., 2011). Consequently, to form a comprehensive, and arguably more sensitive profile of outcomes in tumour populations, it is necessary to administer assessments sometime after surgery when these acute effects have diminished. Indeed, a number of studies report that tumour patients are likely to be in a more stable phase of their recovery at least three months post-operatively, when some functional reorganisation and neuroplasticity has begun (Shafi & Carozza, 2012). Accordingly, follow-up assessments can be used to draw

more valid conclusions and a greater understanding of finer-grained linguistic functions, and the extent to which neurological tumours and surgical resection impact upon language operations.

In addition, in its current form, the BLAST is not intended to address theoretically-motivated questions regarding normal linguistic function and the structures that support it. Specifically, the tasks utilised in the BLAST focus on the single-word level; however, uncertainty remains regarding the extent to which more complex aspects of language, such as sentence-level production and comprehension, can in fact be decomposed into core, cognitively motivated skills. Further, the BLAST does not attempt to distinguish between language-specific processes and other types of cognitive processes that may be required to perform successfully on the tasks. For example, some of the subtasks included in the BLAST protocol may place demands on more general cognitive control processes that are not necessarily specific to language. In the current investigation, we explore these issues in greater detail.

Long-term Language Outcomes in Tumour Surgery Patients

Recent estimates suggest as many as fifty percent of individuals undergoing surgery for left-hemispheric lesions will experience post-operative language and cognitive disruption (Davie et al., 2009). Such disruption has been associated with a detrimental impact upon quality of life and return to vocational activities (Gulati et al., 2009; Heimans & Taphoorn, 2002; Pelletier, Verhoef, Khatri, & Hagen, 2002). Therefore, an important goal of research is to gain information about patients' long-term outcomes with respect to language function. This information can have additional benefits. For example, some studies have demonstrated that speech and neurocognitive decline or difficulties following surgery can be an early indicator of survival, disease progression and/ or tumour regrowth (Thomas et al., 1995). Indeed, in an earlier study that investigated 80 patients with malignant glioma, performance on verbal recall and recognition were found to correlate positively with survival (Meyers, Hess, Yung, & Levin, 2000; see also for similar findings; Brown et al., 2006).

In addition to providing general information regarding long-term outcomes, follow-up studies can also be used to determine factors that predict long-term language capacity. This type of information can be enormously helpful in evaluating cost-benefit ratios associated with particular types of surgical interventions, which are clinically relevant for both the patient and surgeon. Ultimately, a greater understanding of long-term outcomes, consequences, and factors that may influence prognosis can be invaluable for determining the most appropriate course of treatment, as well as guiding criteria for clinical decision making; this ultimately can enable both the patient and surgeon greater capacity to make informed decisions regarding outcomes and prognosis (Armstrong, Goldstein, Shera, Ledakis, & Tallent, 2003; Giovagnoli, Silvani, Colombo, & Boiardi, 2005; Papagno et al., 2012). Moreover, the information obtained from follow-up studies and assessments is necessary to guide more tailored rehabilitation strategies, which is particularly important given the detrimental effects language and cognitive deficits can have on vocational capacity, quality of life, and inter-personal relationships (Le Dorze & Brassard, 1995; Moritz-Gasser, Herbet, Maldonado and Duffau, 2012; Pelletier et al., 2002; Taphoorn & Klein, 2004).

Finally, follow-up studies can highlight the need for appropriate neurocognitive rehabilitation within tumour populations (Giovagnoli, 2012). Specifically, within the New Zealand context, cognitive and linguistic rehabilitation programmes are largely tailored for those affected by vascular lesions and traumatic brain injury. Such observations are attributed to the widespread assumption that persisting difficulties will be insufficiently common to

justify further therapy in neurological tumour patients. Moreover, given metastatic and high-grade neurological tumours are typically associated with poor prognosis (Deorah, Lynch, Sibernaller, & Ryken, 2006), efforts have primarily focused on the identification of treatments to control tumour growth and prolong survival rather than rehabilitation strategies (Gehring, Sitskoorn, Aaronson, & Taphoorn, 2008). Importantly however, without detailed and comprehensive information regarding the incidence and nature of persisting language impairments in tumour populations, this assumption has not been adequately explored using a cognitively-motivated assessment battery tailored specifically for this population.

Studies examining long-term language outcomes in tumour patients

Despite the important clinical implications, relatively little attention has been paid to long-term language outcomes following tumour resection surgery (Finch & Copland, 2014; Wu et al., 2011). Further, those studies that do exist vary enormously in the characteristics of the patient sample (e.g. low grade vs high grade tumours; undifferentiated vs site-specific lesion localisations), the nature of the testing protocol (ranging from formal aphasia assessments to core skills measures and self-report), the time of follow-up testing (ranging from 21 days to nine years' post-surgery), and the definition of 'abnormal' (e.g. variability in cut-off scores and normative data used to identify impairment) (Gehring, et al 2008). Indeed, a recent systematic review highlighted such difficulties when language functions were investigated in a neurological tumour population. Specifically, the majority of potential studies were excluded given they were largely based on single-case studies, reported only pre-surgical or post-surgical outcomes, and included bi-lingual individuals, or those with specific language or cognitive impairments (Finch & Copland, 2014). The following sections will review relevant studies that have evaluated long-term language outcomes.

A number of studies report that language function in tumour patients can improve substantially following tumour surgery, in some cases to normal levels. Recently, Finch and Copland (2014) reviewed nine studies that had examined language function in tumour patients both preoperatively and postoperatively using standardised aphasia assessments. Although the patient populations and methods of assessment varied widely across studies, the majority of studies reported an initial post-surgical deterioration that greatly improved, and in some cases, even resolved three months post-operatively. Several of the most important of studies reviewed by Finch and Copland are summarised in Table 2.3.

Turning now to other studies not included in the Finch and Copland (2014) review,

the most important of these are also summarised in Table 2.3. One such major study was that of Sanai and colleagues (2008). These researchers examined 250 undifferentiated tumour patients immediately pre-operatively, post-operatively, and at least three months following surgery on a protocol that included the following tasks: counting, naming objects, single-word reading, repetition of complex sentences, and single-word and sentence writing. Performance on these tasks was used to identify the following types of language deficits: anomia (defined as an inability to name objects, but intact fluent speech and repetition), alexia (an inability to read or spell words), expressive aphasia (impaired speech or writing), receptive aphasia (fluent meaningless speech accompanied by spoken language comprehension difficulties) and mixed aphasia. Overall, 36 percent of patients demonstrated one of these language deficits preoperatively, and 8.4 percent of patients exhibited a significant deterioration in their language performance one-week post-surgery. Importantly however, these deficits improved three months' post-surgery, with just 2.6 percent of patients showing poorer language functioning relative to their baseline scores at this timepoint, and no patients demonstrating new impairments. Such patterns continued to persist six months' post-surgery, wherein no new deficits were evident, with only 1.6 percent of patients continuing to exhibit permanent language deficits. Taken together, Sanai and colleagues (2008) concluded that language function in tumour patients either improves to baseline performance by three months' post-surgery or not at all.

The above studies suggest a good prognosis for those with language deficits as a consequence of tumour. However, some studies paint a less optimistic picture. Several of the most important of these are summarised in Table 2.3. For example, Satoer and colleagues (2013) assessed the spontaneous speech of 27 patients with left hemisphere glioma (High-grade glioma: HGG: 44.4%, Low-grade glioma: LGG: 55.6%) at two time-points: one month prior to tumour resection and three months post-operatively. Spontaneous speech was assessed using the linguistic computerised programme CLAN (MacWhinney, 1991), which analysed the following qualitative features of speech: *self-corrections, repetitions, lexical diversity, incomplete sentences,* and *mean length of utterance of words (MLUw)*. Participants also completed the Boston Naming Test and a category fluency task. Impairment was defined as scores significantly lower than that of healthy control participants. On the spontaneous speech measures, five of the ten patients who demonstrated difficulties preoperatively continued to demonstrate less fluent speech, as defined by reduced MLUw and greater incomplete sentences. The remaining five patients recovered to normal levels of performance

three months post-operatively. Importantly, seven patients with tumours in eloquent language areas developed new post-operative difficulties on spontaneous speech, most of whom had low-grade glioma (Satoer, Vincent, Smits, Dirvan, & Visch-Brink, 2013).

Similar findings were obtained by Ilmberger and colleagues (2008), who assessed a sample of 153 awake craniotomy patients using the Aachen Aphasia Test (AAT), which classifies individuals according to the classical aphasia subtypes (e.g. global, Broca's, Wernicke's, and amnestic) (Huber et al., 1984). Tumours were localised to several areas, including frontal, frontotemporal, parietal, temporal, insula, and temporoparietal. Overall, 32 percent of patients exhibited a new language impairment three weeks following tumour resection surgery, which was not present during the acute pre-operative phase. In around a fifth of these cases, these impairments were still evident at seven months follow up. Multivariate logistic regression revealed that increased age (> 40 years) and pre-operative aphasia were significant risk factors for persistent aphasic disturbance, as measured on the AAT (Illmberger et al., 2008). Importantly, these patients were assessed using a conventional aphasia assessment; arguably, higher rates of impairment may have been detected using a more sensitive protocol, based on current cognitively-motivated theories of language.

Finally, Dijkstra and colleagues (2009) examined the nature and extent of long-term neurocognitive deficits in a sample of 89 undifferentiated patients with WHO grade I meningioma. All patients had previously received surgical intervention (either complete or partial resection) and were assessed on average 3.4 years following surgery. Their assessment calculated a range of cognitive domains. Language tasks included the category fluency task, the Stroop task, the digit span task and various other verbal memory tasks. Despite the prolonged period since surgery, when compared to healthy control participants, those with meningioma continued to exhibit impairment on the following tasks: Letter-Digit Modalities Test, Rey's Auditory Verbal Learning Test, Working Memory Test, Digit Span, Category Verbal Fluency Test, Stroop Colour Word Interference Test, and the Concept Shifting Test. Moreover, those with left-sided meningioma performed significantly worse than right-sided patients on tasks that made up the verbal memory domain - Rey's Auditory Verbal Learning Test and the Category Verbal Fluency Test. Importantly however, Dijkstra and colleagues (2009) considered impairment as any Z score less than 1.5 standard deviation below healthy control participants, which is a fairly liberal definition compared to the more standard -2/-2.5 cut-off. Whilst this retrospective study is unable to provide a comprehensive evaluation of long-term language functioning due to the absence of pre-operative measures, the findings do

suggest that cognitive difficulties may continue to persist in a sample of meningioma patients.

The probability of long-term language deficits is also likely to be influenced by the location of the tumour/surgical ablation. In a recent study, Papagno and colleagues (2010) examined long-term language outcomes in a sample of 44 individuals with glioma (LGG: 24; HGG:19). All patients had undergone tumour resection surgery three months prior to assessment. Language was assessed on a range of tasks, including *verbal fluency*, *naming of famous faces*, *naming of pictured objects and actions*, *naming by description*, *word, non-word* and *sentence repetition*, and a range of comprehension tasks. In cases where tissue was resected from the temporal lobe and/or the uncinate fasciculus (a bundle of white matter tracts that connect parts of the limbic system with frontal areas), naming either remained the same or deteriorated at follow-up relative to baseline pre-operative scores (Papagno et al., 2010). Moreover, in cases where tissue was resected from the frontal part of the uncinate fasciculus, naming of famous faces was significantly poorer compared to those who underwent frontal resection without the uncinate fasciculus removal. These findings highlight the important role that localisation, particularly in this case the uncinate fasciculus, has in the retrieval of word form for proper names. Such findings are consistent with the observations of lesion location being an important determinant of language functioning (e.g. Faulkner et al., 2017), whilst also demonstrating that lesion location continues to play a significant role in language functioning many months after surgery.

More recently, Papagno and colleagues (2012) evaluated 226 glioma patients (HGG: 11; LGG: 110; ELGG: 105) across varying time periods: one week prior to surgery, one week following surgery, and where possible, every three months. Language and cognitive functioning were assessed using the Milano-Bicocca Battery (MIBIB, see above), with impairment defined as performance below that of normal, healthy controls. Overall, despite some general improvement in functioning three months' post-surgery, a deterioration in performance was evident in some patients. Specifically, during the pre-operative phase, 24.5 percent of left frontal patients showed impaired performance on a letter fluency task, however, during the three-month follow-up, this rate increased to 40 percent. Meanwhile, 41.2 percent of left temporal patients showed pre-operative impairment on this task, which increased to 48.39 percent during the follow-up phase. Similar deterioration was evident on a category fluency task, wherein 25.8 percent of left temporal patients demonstrated impairment three months' post-surgery relative to just 18 percent during the pre-operative phase. Left temporal patients also exhibited long-term decline in a naming task, with 48

percent of patients exhibiting follow-up impairment compared to 41 percent during pre-operative testing.

Taken together, the above findings support a general trend toward favourable long-term language outcomes following tumour resection surgery. However, there is considerable variability in outcomes across studies, which is likely attributable to the disparities in the methodologies used. These include the nature of the assessment protocol (e.g. based on classical model of language, cognitively or linguistically motivated), the nature of the participant sample (e.g. single-case versus group studies, type and grade of tumours), whether the sample was selected specifically for vulnerability to language deficits (e.g. tumours in eloquent cortex versus undifferentiated), and the way in which impairment was defined (e.g. according to aphasia quotients or below normal, healthy control participants). Below we discuss some of the major sources of variability and their possible impact on outcomes.

Table 2.3.

List of studies that have evaluated long-term language outcomes in neurological tumour populations

Outcomes	Authors	Population Sample	Test Battery Administered	Time Points	Long-term Outcomes
Negative outcomes					
	Satoer et al., 2014	One patient, K.O. who presented with a low-grade glioma involving the supplementary motor area	Aachen Aphasia Test (Huber, Poeck, & Willmes, 1984)	One month pre-operatively, two weeks, seven weeks, three months and twelve months post-surgery	12 months post-surgery, generative speech did not reach pre-operative baseline levels.
	Satoer, Vincent, Smits, Dirvan, Visch-Brink, 2013	27 patients with left-hemispheric glioma (HGG: 44.4%, LGG: 55.6%)	Boston Naming Test (Kaplan, Goodglass, & Weintraub, 2001); category fluency task (Luteijn & Barelds, 2004); and spontaneous speech using the linguistic computerised programme, CLAN (MacWhinney, 1991). Speech was assessed for: self-corrections, repetitions, lexical	One month pre-operatively and three months post-surgery	Fifty percent of patients who showed pre-operative impairment recovered, relative to baseline, whilst Fifty percent continued to show less fluent speech and incomplete sentences. Seven patients with tumours in eloquent language, mostly LGG, developed new impairments.

				diversity, incomplete sentences, mean length of utterances.
Ilmberger, Ruge, Kreth, Briegel, Reulen, & Tonn, 2008	153 awake craniotomy patients	Aachen Aphasia Test	Pre-operatively, and three weeks and seven months post-operatively	17.6 percent showed new language disturbances seven months post-surgery, with 10.9 percent of these patients not significantly impaired during the pre-operative period. Age (>40 years) and pre-operative performance were significant risk factors for persistent aphasic disturbance.
Papagno, Casarotti, Gallucci, Riva, & Bello, 2012	226 glioma patients (HGG: 11; LGG: 110; ELGG: 105) (at least one follow-up was collected for 117 patients)	Milano-Bicocca Battery (impairment defined as performance below normal, healthy controls)	One week pre-operatively, one week and three months post-surgery	Whilst a general improvement was evident, pre-operatively, 24.5 percent of left-frontal patients showed pre-surgical impairment on a letter fluency task, which increased to 40 percent at follow-up, whilst 48.39 percent of left temporal patients showed long-term impairment cf to 41.2 percent pre-operatively. On a category fluency task, 25.8 of left temporal patients

					showed long-term impairment, cf 18 percent during the pre-operative phase. A long-term decline was also evident on a picture naming task (48 percent cf 41 percent).
	Dijkstra et al., 2009	89 patients with WHO grade I meningioma.	Letter-Digit Modalities Test; Rey's Auditory Verbal Learning Test; Working Memory Test; digit-span; category verbal fluency, Stroop Colour-word Interference test; and concept Shifting test (Benton, 1968; Houx & Jolles, 1994; Lezak, 2004; Wechsler, 1997). (impairment defined as performance below normal, healthy controls)	On average, 3.4 years following surgical resection (either complete or partial)	Meningioma patients showed significantly lower performance, relative to healthy controls, on all domains, with the exception of attentional functioning.
Positive outcomes					
	Duffau, Gatignol, Mandonnet, Capelle, & Taillandier, 2008	115 patients with LGG (WHO grade 11) undergoing direct electrical stimulation to map lesions in the	BDAE (Goodglass et al., 2001)	Pre-operatively and three months post-surgery	Overall, 98 percent of patients returned to their pre-operative baseline levels within three months following surgery.

47

				eloquent cortex
Teixidor, Gatignol, Leroy, Masuet-Aumatell, Capelle, & Duffau, 2007	Eight LGG patients	Subtests from the BDAE (French adaption, Mazaux & Orgogozo, 1982) and Picture Naming (Metz-lutz, Kremin, Deloche, Hannequin, Ferrand, & Perrier et al., 1991)	Pre-operatively, and three months post-surgery	Five patients recovered their pre-operative verbal working memory scores, and the remaining three significantly improved. Most sensitive to follow-up improvement: BDAE subtests of reading, abstract phase repetition, and sentence dictation.
Sanai, Mirzadeh, & Berger, 2008	250 undifferentiated tumour patients	Counting, object naming, single-word reading, repetition of complex sentences, and single-word and sentence writing. (impairment defined as meeting criterion for anomia, alexia, and expressive, receptive and mixed aphasia)	Pre-operatively, post-operatively, three, and six months following surgery	One month post-surgery, 6.4 percent of patients with pre-operative impairment continued to show language deficits, with just 3.2 percent showing new impairments. Three months post-surgery, just 2.6 percent of patients showed impairment, with no patients showing new deficits. Six months post-operatively, just 1.6 percent of patients were impaired, with no patients demonstrating new impairment.
Santini, Talacchi, Squintani,	22 patients undergoing awake resection for left-	Included letter fluency and subtests from the	Pre-operatively,	Despite an immediate post-operative worsening, the

Casagrande, Capasso, & Miceli, 2012	sided glioma (HGG: 36%, LGG: 64%) (11 patients assessed at follow-up)	Battery for Aphasia Analysis of Deficit (BADA; Miceli, Laudanna, Burani, & Capasso, 1994)	post-operatively, and three-six months follow-up	majority of patients showed a general improvement between the pre-operative and follow-up period.

Tumour type. Converging evidence supports a strong association between the role of tumour grade and long-term language outcomes. Specifically, it is relatively well established in the literature that higher-grade tumours, which have greater infiltration potential, are associated with relatively more impairment severity relative to lower-grade tumours (Heimans & Reijneveld, 2012; Imperato Paleologos, & Vick, 1990; Klein & Heimans, 2004; Whittle et al., 1998). Indeed, in an earlier study evaluating the effect of tumour grade on cognitive functioning, Hom and Reitan (1984) compared patients with high grade (III-IV; n=46) and low grade (I-II n=46) tumours on the Wechsler Adult Intelligence Scale – third edition (WAIS-III; Wechsler, 1997), and select subtests from the Halstead–Reitan Neuropsychological Test Battery (Reitan & Wolfson, 1993). Overall, those with higher-grade tumours performed more poorly on both measures relative to those with lower grade lesions (Hom & Reitan, 1984). More recently, Noll and colleagues (2014) reported that those with low-to-medium grade glioma (II to III) exhibited a greater frequency of impairments on measures of verbal learning and memory, auditory attention, and executive functions, however, patients with high-grade glioma (IV) demonstrated significantly worse performance, and a broader range of deficits (see Noll et al., 2014). Similar findings have been observed in a number of other studies. For example, in an earlier study of undifferentiated tumour patients, Hahn and colleagues (2003) reported that those with high-grade lesions (n= 31) performed significantly more poorly than patients with low-grade lesions (n= 37) on measures of verbal fluency and Trail Making (Hahn, Dunn, Logue, King, Edwards, & Halperin, 2003).

However, significant language deficits are not limited to those with high-grade tumour. They have also been observed in meningioma – a low-grade tumour according to the WHO classification system. For example, in the aforementioned study, Dijkstra and colleagues (2009) examined a sample of 89 meningioma patients, on average 3.4 years following tumour resection surgery. The authors reported significantly lower scores on executive functions, relative to healthy controls. Such findings suggest those with low-grade meningioma are capable of demonstrating long-term neurocognitive deficits following tumour resection surgery. Importantly however, more thorough interpretation is limited given Dijkstra and colleagues (2009) did not directly compare this group with a high-grade tumour group, nor did they conduct pre-operative assessments from which to assess baseline functioning.

Of note, not all studies have reported significant effects of tumour type. For example,

in an earlier study, Scheibel, Meyers and Levin (1996) compared the performance of undifferentiated patients with WHO IV glioblastoma (n= 106) and that of patients with less malignant tumours (< WHO III; n=139). They found no group differences on the WAIS-R measures of verbal IQ, and the digit symbol subtest, nor did they show any differences in performance on the naming and comprehension language subtests of the Multilingual Aphasia Examination (Benton, & Hamsher, 1983); nor on a standard letter fluency task.

Tumour location. Of course, the locus of the tumour is also likely to influence long –term outcomes. Specifically, in the aforementioned study by Dijkstra and colleagues (2009), patients with left-hemispheric meningioma performed significantly worse on measures of verbal memory at long-term follow-up, relative to those with right-hemispheric lesions. More recently, Papagno and colleagues (2012) found lesion location was a significant predictor of relapse across the acute surgical phase to at least three months' follow-up. For example, preliminary results showed performance on delayed verbal recall, face naming, object naming, and verbal fluency was associated with lower scores in left temporal tumour patients. (For similar findings see also Faulkner et al., 2017; Scheibel et al., 1996).

Participant Age. A well-established finding in the literature is that older patients (e.g. > 55 years) typically demonstrate disproportionately greater language difficulties relative to their younger counterparts (Ilmberger et al., 2008; Ostrom et al., 2016; Thomas et al., 1995). Such effects have been found to significantly predict long-term outcomes, when compared to preoperative performance. For example, Ilmberger and colleagues (2008) administered the Aachen Aphasia Test during the preoperative phase to 128 patients who were undergoing awake surgery for tumours near crucial language areas. At seven months' post-surgery, greater age (>40) was found to be a significant predictor of persistent aphasic disturbance. Kaleita and colleagues (2004) also found significant effects of age in a tumour population. For example, 79 patients who were undergoing treatment for undifferentiated malignant brain tumours were assessed on the following measures: Trail Making, Digit Span, and Oral and Written encoding/ decoding of visual processing. A one-way Analysis of Variance (ANOVA) revealed a significant effect of age, with significant differences emerging between the >60 age group and the <30 age group on most individual measures. However, fewer significant differences emerged between the >60 age group and those aged between 36 and 59 (Kaleita et al., 2004).

Other Factors. A number of recent studies have reported a strong association between antiepileptic medication and long-term neurocognitive sequelea (Dijkstra et al., 2009; Eddy,

Rickards, & Cavanna, 2011; Klein et al., 2003). These include reduced attention and concentration, impaired motor performance, lower IQ levels, slower processing speed, impaired memory, and visuomotor functions (Farwell, Lee, Hirtz, Sulzbacher, Ellenberg, & Nelson, 1990; Manni et al., 1993; Riva & Devoti, 1996; Smith et al., 1987). For example, in the aforementioned study involving 89 tumour patients approximately 3.4. years following surgery, Dijkstra and colleagues (2009) reported that current use of antiepileptic drugs (AED) was associated with significantly lower executive function and psychomotor speed relative to patients not currently using AED's (For a full review, see Eddy et al., 2011).

Another huge source of variability is the actual resection itself – that is, whether the surgeons resect the tumour based on conservative or aggressive parameters. This of course varies across cases and surgeons, and also possibly on the date of the study. For example, more recent surgical methods, such as intraoperative MRI and fluorescence-guided surgery, allow for greater supra-maximal resection compared to earlier methods (Bailey & Lucas, 2014), which is likely reflected in differential outcomes across different timepoints. Further, there is growing evidence that post-surgical outcomes are influenced by the extent of surgical resection (EOR; Smith et al., 2008; Potts, Smith, Molinaro, & Berger, 2012). Indeed, in a sample of 158 LGG patients, Smith and colleagues (2008) reported that the EOR was a significant predictor of overall survival, and progression free survival at least two years' post-surgery, even when the effects of age, tumour location and tumour subtype were adjusted Such results led the authors to conclude that a more aggressive surgical resection is a significant predictor of long-term outcomes. Of course, this likely varies depending on the nature of the tumour – those that are smaller, and in non-eloquent areas likely result in greater EOR compared to those that are generally more infiltrative and diffuse. In such cases, many surgeons tend to employ a conservative strategy, wherein the EOR is reduced and a total resection of tumour boundaries is often not possible (Recht, Lew, & Smith, 1992; Reijneveld, Sitskoorn, Klein, Nuyen, & Taphoorn, 2001). Related to this, the degree of tumour margins is also likely to influence long-term outcomes. Specifically, well-defined tumour margins typically produce more favourable follow-up performance, relative to less-defined lesions that are associated with greater compression and infiltration on surrounding tissue (Davie et al., 2009).

Finally, a number of studies have identified that surgical resection of tumours can result in immediate post-operative improvement. For example, in a sample of 29 glioma patients, 24% exhibited immediate postoperative improvement on a number of tests,

including verbal digit span, word fluency, and immediate and delayed recall (Talacchi et al., 2011). Such improvement has been attributed to the reduced intracranial pressure and compression of brain tissue (Heimans & Reijneveld, 2012). Based on these observations, it is likely that tumour volumes play a role in long-term change, namely, the resection of larger tumour volumes would be associated with a greater reduction in intracranial pressure and compression, resulting in greater long-term improvement relative to baseline preoperative scores. In other words, surgical intervention is likely to have a greater impact on follow-up scores for larger tumours due to greater release of intercranial pressure when compared to smaller lesions.

Concluding remarks. To conclude, our review of the literature shows that long-term outcomes for language can vary widely. We identified a number of possible reasons for this wide variation. Our evaluation also raised concerns about some of the assessment methods that are commonly used in this domain. It was argued above that conventional aphasia assessment batteries, which are validated primarily on post-stroke aphasia, might not be sufficiently sensitive at detecting language difficulties in brain tumour patients. Consequently, an assessment better tailored to this population might be more effective at identifying positive – or negative - long-term change. Further, we argued that conventional aphasia assessments – which are primarily based on the classical model of aphasia – might not provide a sufficiently comprehensive picture of the various cognitive processes essential for language, not as they are now understood.

Aims and Predictions of the Present Study

In this study, our overarching objective is to assess long-term language outcomes in a small undifferentiated sample of brain tumour patients using a cognitively-motivated assessment, the BLAST. A particular advantage of this assessment is that it has been shown to be highly sensitive to (preoperative) language deficits in a tumour population (see Faulkner et al., 2017). The BLAST generates scores for a set of eight key cognitive operations believed to be essential for effective language performance (called *core skill scores*). Given patients are considered to be in a more stable phase of recovery at least three months post-operatively, with most language recovery found to occur during this phase (e.g. Finch & Copland, 2014; Sanai et al., 2008; Satoer et al., 2013) our follow-up assessments will be administered at least three months' post-surgery.

There are two primary aims. The first is to examine the nature of, and factors determining change across three phases: immediately pre-operatively, post-operatively, and at least three months' follow-up. Specifically, we wish to:

1. Examine the incidence of significant impairment (core skills scores significantly below those of controls) across the pre-operative and follow-up periods, for the group as a whole.
2. Assess the direction of change across these testing phases for each individual participant on each core skill measure (that is, whether performance improves, declines or stays the same).
3. Assess the extent to which preoperative language performance predicts long-term follow-up performance on each core skill measure.
4. Examine factors that may influence performance change (such as chronological age, tumour volume, tumour type and grade)

The second aim of this study is focused not on the degree of change, but rather to examine the relationship between lesion location and our key language measures at specific timepoints (core skill scores). To do this, patients will be assigned into one of five broad groups based upon tumour localisation (left frontal, left temporal left parietal, right frontal, and right parietal). We will then perform group comparisons to determine whether these groups show reliably different profiles. Our specific hypotheses (based on Faulkner et al., 2017) are summarised below and operationalised in Table 2.5.

Predictions. We chose to base our predictions for long-term change from Papagno and colleague's (2012) findings. The reasons for this were two-fold: i) their assessment battery was distinct from previous studies as it was tailored specifically for neurological tumour patients; and ii) their protocol was cognitively-motivated, encompassing tasks that place emphasis on the contributions of higher-level skills. Based on this study, our predictions were as follows:

a. At long-term follow-up, scores are not expected to reach the same levels as baseline preoperative performance (as assessed in T scores).

b. Pre-operative scores for a particular core skill will be a significant predictor of follow-up scores on subsequent tests.

c. Any performance changes that are documented will be dependent on the following factors: age, lesion volume, and tumour grade and type. Our predictions and operationalisations regarding these factors are summarised in Table 2.4.

Table 2.4.

Factors predicted to modulate change in scores from preoperative to follow-up testing phases.

	Prediction and operationalisation
Chronological Age	Greater age will be predictive/ associated with lower general improvement and/or greater deterioration on the BLAST core skill measures, as measured using logistical regression (age will serve as the continuous dependant variable, whilst the absolute change in scores from pre-operative to follow-up period will serve as the dependant variable)
Lesion Volume (cm^3)	Smaller lesion volume (cm^3) will be predictive/ associated with lower overall improvement and/or greater deterioration in core skill scores, as measured using logistical regression (lesion volume (cm^3) will serve as the continuous dependant variable, whilst the absolute change in scores from pre-operative to follow-up period will serve as the dependant variable)
Tumour grade and type	Higher-grade tumours and specific tumour types (WHO III-IV; glioblastoma, high-grade glioma, and astrocytoma), will be significantly predictive/associated with lower improvement and/or greater general deterioration in core skill scores, when compared to low grade tumours (WHO Grade I), as measured using logistical regression (tumour grade will serve as the dependant variable, whilst the absolute change in scores from the pre-operative to follow-up period will serve as the dependant variable)

For the second aim of our study, which was to assess the effect of lesion location on language profiles at specific timepoints, our predictions were based upon those of Faulkner and colleagues (2017). These are summarised in Table 2.5.

Table 2.5

Anatomical predictions for each core skill measure, as outlined in Faulkner et al (2017).

Cognitive Skill	Implicated Brain Region	Tumour Group Predicted to exhibit lowest average score
Accessing Semantic Knowledge	Left anterior temporal regions	Left temporal patients
Lexical Selection	Left posterior temporal region	Left temporal patients
Auditory Word Recognition	Left posterior superior temporal lobe	Left temporal patients
Goal-Driven Response Selection	Left inferior frontal gyrus	Left frontal patients
Verb Retrieval	Left ventrolateral gyrus	Left frontal patients
Articulatory Motor Programming	Left inferior frontal regions, including the left insula	Left frontal patients
Verbal Short-Term Memory	Left inferior parietal cortex and middle temporal gyrus[2]	Left temporal and parietal patients
Phonological Encoding	Left superior temporal and left parietal lobe[2]	Left temporal and parietal patients

---

[2] Critical regions impinge on both temporal and parietal areas (see Table 2.2.)

2.1. Participants

2.1.1 Tumour patients.

In total, 28 patients admitted for brain tumour surgery at Wellington Hospital in New Zealand, participated in this study from May 2013 until August 2016[3]. All had been assessed immediately pre and post-operatively on the BLAST protocol. The original selection criteria for the current study were broad: Eligible patients were identified by the participating neurosurgeon, Mr Andrew Parker. All patients undergoing craniotomy for debulking or complete resection of a cerebral tumour were eligible, irrespective of aetiology, location, and malignancy (subcortical tumours were excluded). Exclusionary criteria were as follows: i) the patient was under 18 years of age; ii) the patient was deemed not capable of giving informed consent; iii) the patient had any prominent visual or motor coordination disturbances; and iv) English was not their native language. No participants were excluded due to prominent language, visual or motor changes during the pre-operative phase. For our follow-up assessments, patients were excluded if they were currently undergoing chemotherapy or radiation therapy, however, in these instances, they were contacted once they had completed their course of treatment.

Of the 59 patients who were eligible to participate in the follow-up study (pre and postoperative assessments had been completed by the current author, and as part of a previous study, see section 2.4.), 22 passed away before assessment could occur, five could not be contacted, and four were not asked due to unforeseen medical complications, such as hemi-paralysis or loss of vision. Consequently, 28 patients participated in the current follow-up study. However, only 25 patients completed pre and post-operative testing, whilst the remaining three were unable due to time constraints during the acute surgical period; these three patients (S.N., W.R., & D.P.) are therefore not included in the first investigation, which examines pre and post-operative performance, however, they are included in the remaining investigations, which exclusively examine follow-up performance. Table 2.6 provides demographic and general medical information about each patient (See Appendix A for a brief case description and MRI scan of each patient tested in the current study). As can be seen from the Table, the sample includes individuals with a range of tumour diagnoses, including: meningioma, low-grade glioma, high-grade glioma, glioblastoma, and astrocytoma. These

---

[3] Only two patients, WR and DP were assessed in 2016 whilst the majority were assessed between 2013-2015

diagnoses were inferred from medical and radiography reports written by the patient's primary neurosurgeon and radiologist. Based on the histological features of their tumours, patients were categorised into one of two broad groups: *high-grade* (n= 5) and *low-grade* (n= 23). Patients were also categorised into one of five more specific groups, based on the type of tumour they presented with: *meningioma* (17), *low-grade glioma* (3), *glioblastoma* (3), *astrocytoma* (3) *high-grade glioma* (2).

The time since surgery ranged from three to 12.2 months, with the average being 5.7 months. Prior to the preoperative testing, none of the patients had undergone any resection/ debulking surgery. Six patients (S.O, A.E.K., L.W., S.H., G.M., and N.O.H) underwent chemotherapy and/or radiation therapy following their surgery[4]; therefore, to avoid potential confounds of treatment, follow-up testing was commenced between two to four months after treatment was complete (see Table 2.6). In addition, at the time of follow-up testing, recent MRI scans had ruled out further tumour regrowth in 11 patients (A.V.G., J.B., L.W., R.F., S.O., S.G., W.R., B.P., S.N., K.G., and K.W.), however this information was not available for the remaining patients.

Prior to follow-up testing, patients were asked if they had developed any physical weakness, visual problems, impaired motor control, or hearing impairments that may affect their performance on the tasks; however, no difficulties were reported in the current sample. Patients were also asked whether they were currently taking anti-epileptic medication; this information was recorded and presented in Table 2.6. All patients responded with their dominant hand (2/28 were left-handed: K.W. and S.G.). This study received approval by the Capital and Coast District Health Board's Ethics Committee.

---

[4] Five patients had glioblastoma or high-grade glioma type tumours, whilst L.W, who had a right frontal meningioma, was diagnosed with non-metastatic lung cancer following her tumour resection and therefore completed one round of chemotherapy.

Table 2.6.

The relevant demographical and clinical information for each patient, including chronological age, tumour type and malignancy, tumour location and volume, surgical procedure, antiepileptic medication at the time of follow-up, and whether the patient underwent chemotherapy or radiation. (N.B. MRI scans were not available for three patients: SN, SG, PC).

Patient	Age	Gender	Tumour Specimen	Anatomical Group	Tumour Volume (cm^3)	Anti-Epileptic Medication	Malignancy	Surgery Type	Chemotherapy Radiation
51+ Age group									
RF	65	M	Meningioma	Left Frontal	17.5	No	Low	Craniotomy & removal	No
SO	58	M	Glioblastoma	Left Parietal	15.02	No	High	Craniotomy & debulking	Yes
CR	75	F	Meningioma	Left Parietal	51.57	No	Low	Craniotomy & removal	No
BD	62	F	Meningioma	Left Parietal	35.97	No	Low	Craniotomy & resection	No
LW	66	F	Meningioma	Right Frontal	6.4	Yes	Low	Craniotomy & resection	Yes

AEK	52	F	High-grade Glioma	Right Posterior	39.66	Yes	High	Craniotomy & resection	Yes
GP	73	M	Meningioma	Right Frontal	56.98	Yes	Low	Craniotomy & resection	No
SH	55	M	Glioblastoma	Left Temporal	44.49	Yes	High	Craniotomy & resection	Yes
NOH	76	F	High-grade Glioma	Left Parietal	7.26	Yes	High	Craniotomy & resection	Yes
BS	56	F	Meningioma	Right Frontal	30.04	No	Low	Craniotomy & resection	No
GM	70	M	Glioblastoma	Left Temporal	36.5	Yes	High	Craniotomy & resection	Yes
SN	74	M	Meningioma	Right Frontal	na	No	Low	Craniotomy & resection	No
RS	66	M	Meningioma	Right Frontal	4.61	No	Low	Craniotomy & resection	No
WR	60	M	Meningioma	Right Frontal	4.3	No	Low	Endoscopic resection	No

31-50 age group

CA	47	F	Meningioma	Left Temporal	83.57	No	Low	Craniotomy & resection	No

AVG	42	F	Meningioma	Left Parietal	71.33	No	Low	Craniotomy & resection	No
JB	47	F	Meningioma	Right Posterior	41.11	No	Low	Craniotomy & resection	No
AE	46	M	Meningioma	Left Frontal & Right Frontal	na	Yes	Low	Craniotomy & removal of left frontal lesion	No
TD	30	M	Low-grade Glioma	Left Frontal	151.59	No	Low	Craniotomy & debulking	No
DF	40	M	Low-grade Glioma	Right Frontal	50.36	No	Low	Craniotomy & removal	No
NH	42	F	Meningioma	Right Frontal	na	No	Low	Craniotomy & resection	No
SG	42	M	Astrocytoma (LG)	Right Posterior	49.03	Yes	Low	Craniotomy & resection	No
KG	41	F	Low-grade Glioma	Left Parietal	15.76	Yes	Low	Awake craniotomy & resection	No
PC	34	M	Meningioma	Right Posterior	61.28	No	Low	Craniotomy & resection	No

DP	30	M	Astrocytoma (LG)	Left Frontal	6	No	Low	Awake craniotomy & debulking	No
LC	43	F	Astrocytoma (LG)	Left Frontal	46.75	No	Low	Craniotomy & resection	No

18-30 age group

BP	18	F	Meningioma	Left Frontal	15.56	Yes	Low	Craniotomy & resection	No
KW	21	M	Meningioma	Left Temporal	8.77	Yes	Low	Craniotomy & resection	No

Each patient's pre-surgical MRI scans and radiology reports were used to provide anatomical descriptions of tumour localisation. For each patient that had an available preoperative MRI scan, lesion location was categorised according to the region containing the largest percentage of lesioned voxels (this was calculated by first normalising the preoperative scan using MRIcron; Rorden, Karnath & Bonilha, 2007). This information determined the categorisation of patients into one of the following five broad anatomical groups: *left frontal, left temporal, left parietal, right frontal, right posterior* (see Table 2.7)[5]. For those patients where there was no preoperative MRI scan available, lesion location was determined by preoperative radiology and surgical reports. Left and right lateralised tumours were distinguished by the longitudinal fissure, whilst frontal and parietal/posterior tumours were distinguished by the central sulcus boundary. Two patients (S.O. and K.G.) had tumour boundaries that crossed the central sulcus. In these cases, patients were categorised according to the region that contained the largest percentage of lesioned voxels. Table 2.7. displays the total number of patients per anatomical group. One patient with multiple tumours (A.E.) was not assigned to any anatomical category, and was not included in any anatomical analysis.

Table 2.7.

The total number of patients per anatomical categorisation.

Anatomical Group	Follow-up Patients
Left Frontal	5
Left Temporal	4
Left Parietal	6
Right Frontal	8
Right Posterior	4
Not assigned to a category[6]	1

---

[5] MRI scans were not available for three patients: SN, SG, & PC. Therefore, in these cases, tumour localisation was determined from the MRI and surgical reports

[6] Patient AE who presented with multiple lesions, with only one resected (left frontal) at the time of follow-up. For this reason, AE was not assigned to any anatomical category and was excluded from any group analysis

Each patient's pre-surgical radiology reports were also used to provide lesion volumes (cm^3) of each patient's tumour (see Table 2.6.).

Figure 2.2. presents a lesion overlap map for the entire patient sample (n= 24), whilst Figures 2.3 to 2.7 present lesion overlap map for each of the five tumour localisation groups. These overlap maps were created by first manually segmenting each patient's lesion map onto a T1-weighted structural image of their brain using MRIcron (Rorden et al., 2007). Following this, the maps were then normalised by unified segmentation normalization algorithm from Rorden, Bonilha, Fridriksson, Bender and Karnath (2012). Generally, there appears to be greater coverage of the left hemisphere, particularly in anterior regions. Cooler colours (e.g. blue and green) represent less overlap, whilst warmer colours (e.g. yellow and orange) represent more overlap. Figure 2.2. also shows that four patients had lesion overlap in the right frontal region (as indicated by orange), whilst four patients also had some overlap in left temporal regions. Right posterior regions appear to have the sparsest coverage of all.
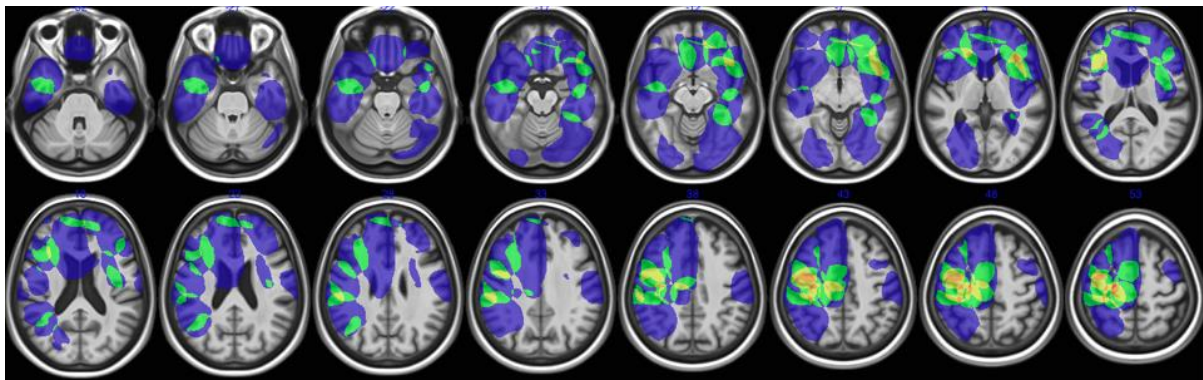
Figure 2.2. Lesion overlap map for all tumour patients[7] showing axial slices on a standard Template (Rorden et al. 2012) at MNI Z coordinates: -32, -27, -22, -17, -12, -7, 3, 13, 18, 28, 33, 38, 43, 48, 53. Red = overlap between five individuals; Orange = overlap between four individuals, yellow = overlap between three individuals; green = overlap between two individuals, blue = no overlap, lesion is specific to only one individual. This map was created from T1-weighted structural images collected at 1.5 Tesla. The lesions were segmented manually onto a 1x1x1mm T1-weighted structural image using MRIcron (Rorden et al., 2007), consulting the T2-weighted FLAIR image for additional guidance. Subsequently, the scans and the lesion maps were normalised in SPM8 (Ashburner et al., 2012) using the unified segmentation normalization algorithm from Rorden et al. (2012) and then superimposed onto a standard template.

---

[7] Three patients did not have brain scans, and were therefore excluded from this analysis
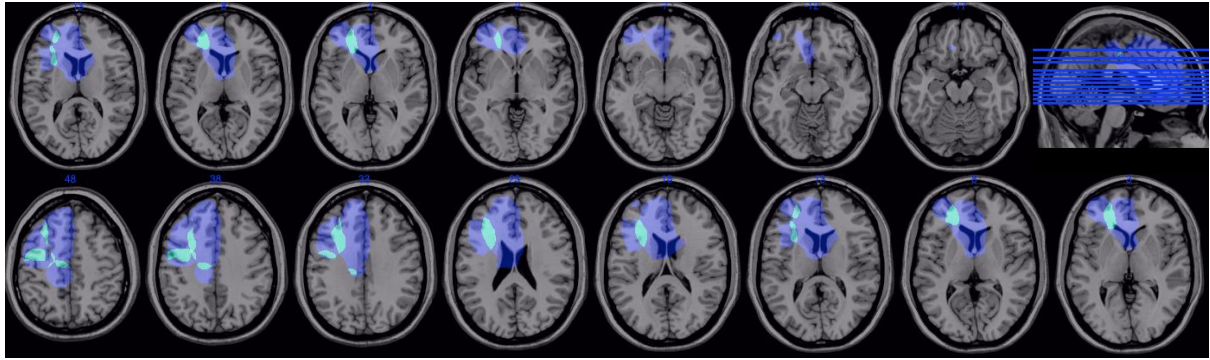
Figure 2.3. Lesion overlap map for the left frontal tumour group. Lesion overlay map for individuals in the left frontal group (N=6). Region showing axial slices on a standard template (Rorden, Bonilha, Fridriksson, Bender & Karnath, 2012) at MNI Coordinates: 48, 38, 33, 23, 18, 13, 8, 3, -2, -7, -12, -17. Red = overlap between five individuals; orange = overlap between four individuals; yellow = overlap between three individuals; green = overlap between two individuals; blue = no overlap, i.e., lesion is confined to one individual.
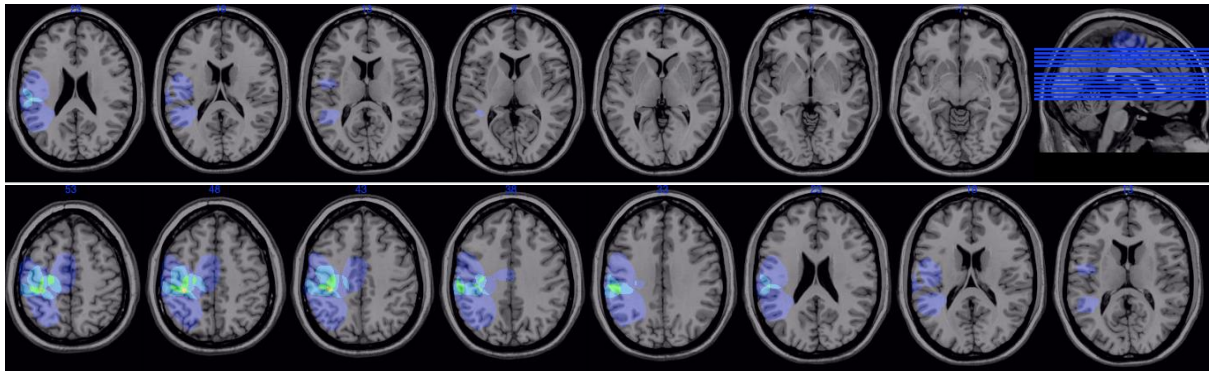
Figure 2.4. Lesion overlap map for the left parietal tumour group. Lesion overlay map for individuals in the left frontal group (N=6). Region showing axial slices on a standard template (Rorden, Bonilha, Fridriksson, Bender & Karnath, 2012) at MNI Coordinates: 53, 48, 38, 33, 23, 18, 13, 8, 3, -2, -7, -12, -17. Red = overlap between five individuals; orange = overlap between four individuals; yellow = overlap between three individuals; green = overlap between two individuals; blue = no overlap, i.e., lesion is confined to one individual.
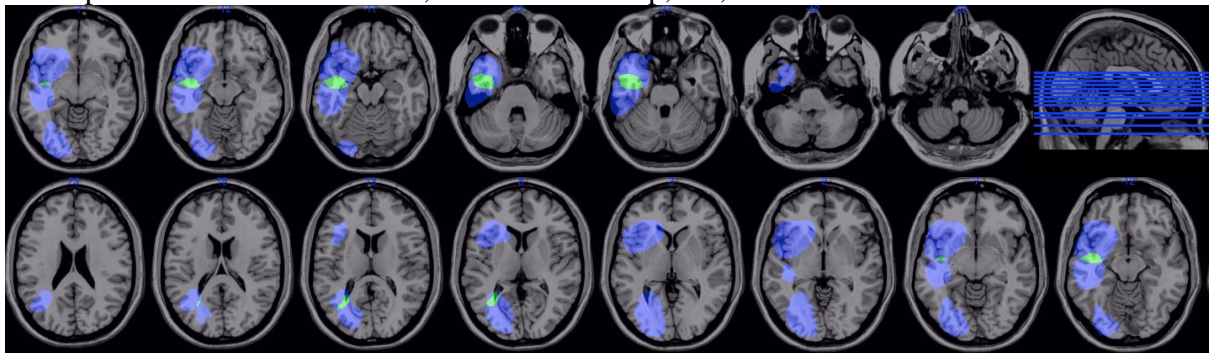
Figure 2.5. Lesion overlap map for the left temporal tumour group. Lesion overlay map for individuals in the left frontal group (N=4). Region showing axial slices on a standard template (Rorden, Bonilha, Fridriksson, Bender & Karnath, 2012) at MNI Coordinates: 23, 18, 13, 8, 3, -2, -7, -12, -17, -32, -27, -42, -52. Red = overlap between five individuals; orange = overlap between four individuals; yellow = overlap between three individuals; green = overlap between two individuals; blue = no overlap, i.e., lesion is confined to one individual.
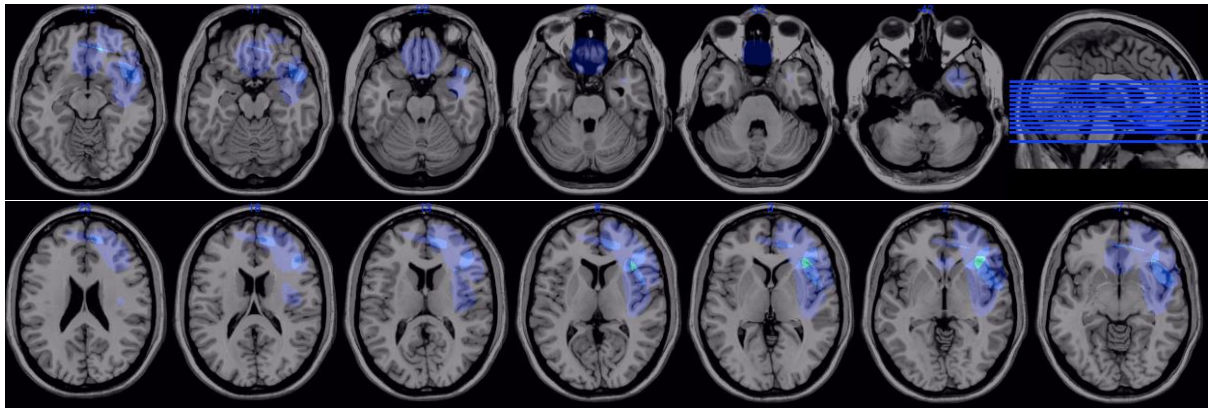
Figure 2.6. Lesion overlap map for the right frontal tumour group. Lesion overlay map for individuals in the left frontal group (N=8). Region showing axial slices on a standard template (Rorden, Bonilha, Fridriksson, Bender & Karnath, 2012) at MNI Coordinates: 23, 18, 13, 8, 3, -2, -7, -12, -17, -27, -32, -42. Red = overlap between five individuals; orange = overlap between four individuals; yellow = overlap between three individuals; green = overlap between two individuals; blue = no overlap, i.e., lesion is confined to one individual.
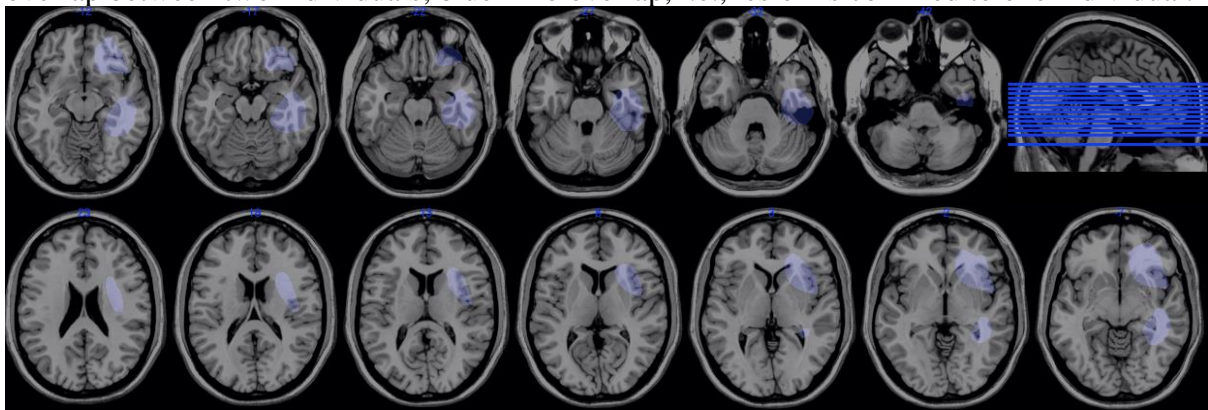
Figure 2.7. Lesion overlap map for the right parietal tumour group. Lesion overlay map for individuals in the left frontal group (N=4). Region showing axial slices on a standard template (Rorden, Bonilha, Fridriksson, Bender & Karnath, 2012) at MNI Coordinates: 23, 18, 13, 8, 3, -2, -7, -12, -17, -27, -32, -42. Red = overlap between five individuals; orange = overlap between four individuals; yellow = overlap between three individuals; green = overlap between two individuals; blue = no overlap, i.e., lesion is confined to one individual.

2.1.2 Healthy Controls

BLAST. 60 neurologically healthy controls had previously completed three versions of the BLAST as part of a previous study (Faulkner, 2015; see also Faulkner et al., 2017). Full details appear in those publications. To summarise, 20 of the participants were aged between 18 and 29 (mean age 22.35; 10 males and 10 females), 20 were aged between 30 and 50 (mean age 37.85; 7 males and 13 females) and 20 were aged 51 years or over (mean age 68.9; 7 males and 13 females). Independent t tests revealed no significant gender-related differences, and all participants spoke English as their native language.

2.2 Design

The current study adopts a case series design. Patients were assessed over three testing phases: one day before surgery (*pre-operatively*), one to three days after surgery (*post-operatively*), and at least three months post-operatively (*follow-up*). During the three phases, patients were assessed on eight subtests from the BLAST protocol (Faulkner et al., 2017). In the follow-up phase, patients also completed a number of additional tasks, which will be discussed in Chapters 3, and 4. The BLAST includes several alternative versions of some subtests. In the preoperative assessment, we administered Version 1 of all tests. In the immediate postoperative assessment, we administered Version 2 of those tests where one existed. In the follow-up assessment, we administered Version 3 of those tests where one existed; otherwise, we administered Version 1. Unless otherwise specified, both response latency and accuracy scores were measured for each task, with individual scores then compared with the relevant age-matched healthy control group (see section 2.3. below).

2.3. BLAST Subtests

The BLAST was administered according to the original protocol outlined by Faulkner et al (2017), and comprises the subtests listed in Table 2.8. A summary of the materials and method for each task are also provided in Table 2.8, and Appendix B gives further details on the procedure for administration, including the various frequency ratios and stimulus items (this information can also be found in Faulkner, 2015 and Faulkner et al., 2017). Subtests were administered in the order listed in Table 2.8[8]. Four tasks were administered orally (Real-word and non-word Repetition, Verbal Fluency, and Articulation), and the remainder (Picture Naming, Verb Generation, Picture-word Verification, and Stroop) were administered on a MacBook Pro laptop using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). During the computer tasks, individuals were situated in front of the laptop, in a position that ensured clear viewing of the screen. For those subtests that contained an audio beep signalling the commencement of a new trial, at least two practice trials were given to ensure the signals were audible. All verbal responses were digitally recorded using Audacity software (Audacity Team, 2008). Responses were also scored according to the procedures in Faulkner et al. (2017), which stipulate that the first response is scored for accuracy, even if subsequently corrected. All experimental items were presented in a fixed pseudo-random order. Where specified, latency data was also recorded using Audacity software (Audacity

---

[8] If a patient was unable to complete the entire testing protocol due to fatigue or time contrasts, the repetition subtests were excluded as these have been found to be the least sensitive in neurological tumour patients compared to the remaining subtests from the BLAST (see Faulkner, 2015)

Team, 2008). In these instances, latency was recorded from the onset of the stimulus tone until the onset of the participant's first response. Filler words such as "um" and "ah" were ignored. If an article (e.g. *the* furniture) was used prior to the target word, response latency was measured from the onset of the tone until the onset of the article/ modifier (see Faulkner, 2015). Prior to data analysis, latency data were trimmed of outliers. To do this, the data was first Winsorised, wherein the longest response latency was replaced by the second longest response latency, and the second longest response latency was replaced by the third longest response latency. Then, any latencies that exceeded two and a half standard deviations from the mean were removed.

Table 2.8.

Summary of the tasks contained in the BLAST, and their associated materials and procedures

Test	Description	Materials	Procedure	Contributes to the following core skills
Picture Naming	Produce the name of a visually displayed object or animal	60 items: 20 depict low frequency nouns, 20 depict medium frequency nouns, and the remaining 20 depict high frequency nouns. Three versions are used - version one (v1) for pre-operative, version two (v2) for post-operative, and version three (v3) for follow-up. Each of these frequency groups contain equal numbers of monosyllabic (20), bisyllabic (20) and polysyllabic (20) items, thus creating nine different frequency x length combinations.	Trials begin with a fixation cross which is replaced after 100ms with the picture stimulus, accompanied by an auditory tone. The picture stimulus remains visible until a response is made. Response scoring comprised of the total number of correct responses (/60).	Accessing Semantic Knowledge, Lexical Selection, Phonological Encoding.
Verb Generation	View a pictured item, and name an action associated with it (e.g., *scissors* -> "cut")	45 items: picturable nouns that act as the stimulus to elicit a verb - 23 depict low-selection items that elicit one specific verb (e.g., heart –> "*beat*"), 22 depict high-selection items that elicit multiple verb responses (e.g., duck –> "*waddle*", "*quack*", "*swim*").	Each picturable item is presented simultaneously with its corresponding auditory word. Participants are instructed *"what the object does, or what is done with that object"*. After three consecutive incorrect responses, feedback is provided from the	Verb Retrieval, Goal-Driven Response Selection.

		This task comprises two alternate versions; v1 is used for pre-operative and follow-up testing, and v2 is used for post-operative testing. In each version, noun stimuli were balanced for name frequency.	examiner. Two practice trials precede the 45 experimental trials. Response scoring comprised of the total number of correct responses (/45).	
Picture-Word Verification	Determine whether an auditory word is the same as a visually displayed item.	48 items: 12 depict pictorial stimuli that are *identical* to the auditory word (e.g. *chair -> chair*); 12 depict a picture that is *phonologically related* to the auditory word, with at least the first two phonemes sharing the target word and containing the same number of syllables and stress pattern (e.g. *chair -> cheque*); 12 depict a picture that is *semantically related* to the auditory word (e.g. *horse* is matched with *deer,* rather than another animal that doesn't share similar semantic features), and 12 depict a picture that is unrelated to the auditory word (e.g. *carrot -> table*). Three alternate versions are used: v1 for preoperative testing, v2 for postoperative testing, and v3 for follow-up testing.	Trials begin with a fixation cross which is replaced after 100ms with the picture stimulus, accompanied simultaneously by the auditory word. Participants are asked to indicate whether the auditory word matches the picture displayed on screen by responding "yes" or "no". Two practice trials precede the experimental trials. Response scoring comprised of the total number of correct responses (/48).	Accessing Semantic Knowledge, Auditory-Word Recognition.

Real-Word Repetition	Immediately repeat an aurally presented real word	60 items from the PALPA word and non-word repetition test (PALPA, Test 9; Kay, Lesser, & Coltheart, 1996). 30 words depict high imageability items (e.g. *potato*) and 30 depict low imageability items/representations (e.g. *valour*). Each of these items are further divided into 15 high frequency words (e.g. *summer*) and 15 low frequency words (e.g. *satire*). One version was administered across all testing phases.	A single word is read aloud by the examiner. Participants are required to immediately repeat the word back. Each word item is presented slowly in a flat intonation, and to avoid lip reading, the experimenter is orientated in such a way as to prevent participants seeing their lips. Response scoring comprised of the total number of correct responses (/60).	Auditory-Word Recognition, Verbal Short-Term Memory.
Non-Word Repetition	Immediately repeat an aurally presented non-word (e.g., slurch)	30 items from the PALPA word and non-word repetition test (PALPA, Test 9; Kay, Lesser, & Coltheart, 1996). Each item differs from a corresponding item in the real-word repetition test by just one letter (e.g. analogy -> *atalogy*). One version was administered across all testing phases.	The procedure is identical to real-word repetition, described above. Response scoring comprised of the total number of correct responses (/30).	Phonological Encoding, Verbal Short-Term Memory.
Stroop task	View a written colour word (e.g., *blue*), and	Adapted from the original Stroop (1935) paradigm. 21 single coloured items from a pool of eight different alternatives (pink,	Trials begin with a fixation cross, which is replaced after 100ms, followed immediately by the target stimulus accompanied simultaneously	Goal-Driven Response Selection.

72

	name the colour it is presented in	black, red, blue, green, orange, yellow and purple). Seven depict congruent items (the word colour matches written name, e.g. *GREEN*). 14 depict incongruent items, with the word colour not consistent with the written name (e.g. *GREEN*). One version was administered across all testing phases.	by a single tone. Each target stimulus is presented in size 60 font. Participants must ignore the identity of the word and instead name the colour that the word is presented in. Two practice trials, both from the incongruent condition, precede the experimental trials. For this task, response latency was also measured, using the procedures outlined in Faulkner (2015) – see section 2.3.	
Letter Fluency	Produce as many words as possible that start with a given letter in 60s	Three letters are given; the first letter is *F*, followed by *A* and then *S* (Spreen, 1998). The instructions were derived from the Controlled Oral Word Association Test (a subtest of the Multilingual Aphasia Examination: MAE; Benton, Hamsher, & Sivan, 1994). Participants are instructed that any words are permitted, with the exception of proper names (for example, "*Boston*" or "*Bob*"), or variations of the same word (e.g. "eat" -> "eating"). A stopwatch is used to time each condition. One version was administered across all testing phases.	Participants are orally provided a letter of the alphabet, and required to produce as many words that start with that letter within 60 seconds. Proper names, variations of the same word, and repetitions are considered incorrect, and are classed as either inappropriate or preservative errors. Data analysis comprised of the total number of correct responses per trial.	Goal-Driven Response Selection, Accessing Semantic Knowledge.

Category Fluency	Produce as many words as possible that belong to a given category in 60s	Identical to letter fluency, with the exception that two categories are given: *animals,* followed by *fruit* (Spreen, 1998).	The Category fluency task adopts an identical procedure to the letter fluency condition, with the exception that participants are asked to generate as many words that belong to a given category (animals and fruit).	Accessing Semantic Knowledge.
Articulatory Agility	Repeat a verbal sequence as many times as possible within five seconds (e.g., *fifty-fifty*)	Based on the Verbal Agility Subtest contained in the Boston Naming Test (BNT; Goodglass et al., 2001). Seven stimulus items are presented in the following fixed order across all testing session: *Mamma, Tip-Top, Fifty-Fifty, Thanks, Huckleberry, Baseball Player, Caterpillar.*	Participants are orally provided with a word or phrase stimulus, and are then required to repeat it as many times as possible within a five-second period. The total number of words correctly repeated for each item within the designated time period is recorded. Responses that are slurred, and/ or incoherent are considered incorrect and excluded from analysis.	Articulatory-Motor Programming.

General procedure

2.4. Pre and post-operative assessments.

All pre and post-operative assessments took place on the neurosurgical ward of Wellington Hospital, New Zealand. Some of these acute assessments were administered as part of a previous project. Specifically, pre and post-operative assessments had been conducted previously for the following 10 patients: R.F., S.O., C.R., B.D., L.W., G.P., C.A., A.V.G., J.B., and A.E. All potential participants were identified by the participating neurosurgeon, Mr Andrew Parker. These patients were then approached by the current author and asked whether they would like to participate in the current study. For all patients, pre-operative testing occurred the day prior to surgery, with potential participants informed that testing would occur over three phases, however they were free to withdraw at any stage. Prior to consent being obtained, patients were given an information and consent form (see Appendix C) and were reminded that their participation was voluntary. All patients were informed that the session would be audiotaped, however their identity would remain confidential. Once informed consent was obtained, patients were asked to comment on any visual problems or motor difficulties. If a patient reported severe difficulties in these domains (for example, blind in one eye, colour blind, or hemi-paralysis), we did not proceed with testing. Patients were also asked to use any aids that may maximise their performance, such as glasses, contact lenses and hearing aids. Testing typically took between 25 – 40 minutes, however this varied depending on individual fatigue levels and discomfort, with individuals reminded that they could initiate breaks at any time.

Post-operative assessments occurred one to three days after surgery. The timing of this session was again variable due to individual differences in fatigue and post-operative discomfort, but typically took between 25-40 minutes. If a patient reported that they felt able to complete post-operative testing, the BLAST administration occurred in the same manner as pre-operative testing. However, where alternate versions of the BLAST were available, version two was administered post-operatively. Once testing was concluded, patients were thanked for their time and asked whether they consented to be contacted in approximately three months' time regarding further follow-up testing. The experimenter reminded each patient that they were not obligated to participate in a follow-up study, and were free to decide in the following months.

All patients who were assessed pre and post-operatively gave their initial consent to be contacted regarding a follow-up assessment. For a full description of the pre and post-operative procedure, please refer to Faulkner (2015).

2.5 Long-term follow-up assessment.

At least three months following surgery, patients were sent a letter detailing the study (see Appendix C), and were then contacted by phone one week later. During this phone call, patients were asked whether they would like to participate in a further follow-up assessment. All patients were informed that testing would take approximately one and a half-to-two hours, and would likely take place in their home, or a place where they felt most comfortable. Patients were reminded that their involvement was voluntary, and were not obligated to participate. All follow-up assessments were conducted by the current author, and were administered in the patient's own home, with the exception of P.C. and R.S., who were tested in office settings at their place of work. The testing procedures were the same as for the preoperative and postoperative session, with the following exceptions. First, patients were asked to note any medication they were currently taking that was directly associated with their tumour, or that they felt may affect their cognitive functioning (for example, anti-seizure, steroids, and chemotherapy drugs). Second, eight additional tasks were also administered which examined other aspects of language and cognition not assessed in the BLAST (see Chapters 3 and 4). These other tests were administered first, followed by the BLAST subtests, which were always administered last, in the order previously outlined. Appendix C provides a table of the entire task administration order during the long-term follow-up. Where possible, we attempted to alternate between computer tasks and tasks of other formats (e.g. oral and pen and paper) to minimise eye strain. This order remained constant for all participants.

Follow-up testing typically took place over one session, however two patients (S.O. and L.W) required two separate sessions over two weeks due to fatigue. Again, patients were reminded that they were free to initiate breaks at any time during the testing and were free to withdraw at any stage. Assessments usually took between one and a half to two hours to complete, however this was largely dependent on individual performance, time constraints, and concentration and fatigue levels

Results

BLAST Accuracy across tasks. Before reporting the main outcome measures for the BLAST – the various core skills scores – we first explored participants' overall accuracy on each of the tasks in the protocol. We recorded the percentage of items correctly responded to for each task, with the exception of the articulation and category and letter fluency tasks. In these cases, the value recorded was the total words produced within a five second period, or the total number of correct items produced within the one-minute time period respectively.

To assess accuracy across tasks, we compared each patient's accuracy score (in percentage) for each task with that of the appropriate aged-matched control group using Crawford and Howell's (1998) modified t-test[9]. This form of comparison was selected as it treats the control sample as statistics rather than parameters, and has been confirmed as a robust method that controls for skewedness in the data, as well as controlling for type 1 error rates, regardless of the size of the control sample. Accordingly, this approach allows for a more stringent method of detecting significant impairment. Moreover, this method has been found to have greater power to detect type II errors compared to other methods (see Crawford & Garthwaite, 2006).

Using the modified t test comparison, impairment was defined as a significant difference in accuracy scores between the individual patient and their age-matched control group, using a p value of $p<.05$. The results are shown in Figure 2.8., which provides a summary of the total percentage of tumour patients who were significantly impaired on each task, relative to their respective control group. Appendix D presents each patient's individual scores for each BLAST subtest.

---

[9] This test is an adapted t test, and uses the following formula: $(t = (x^* - \bar{x})/s\sqrt{(n + 1/n)}$. The x^* is the patient's score; $\bar{x}$ and s are the mean and standard deviation of the control sample respectively; and the n is the size of the control sample.
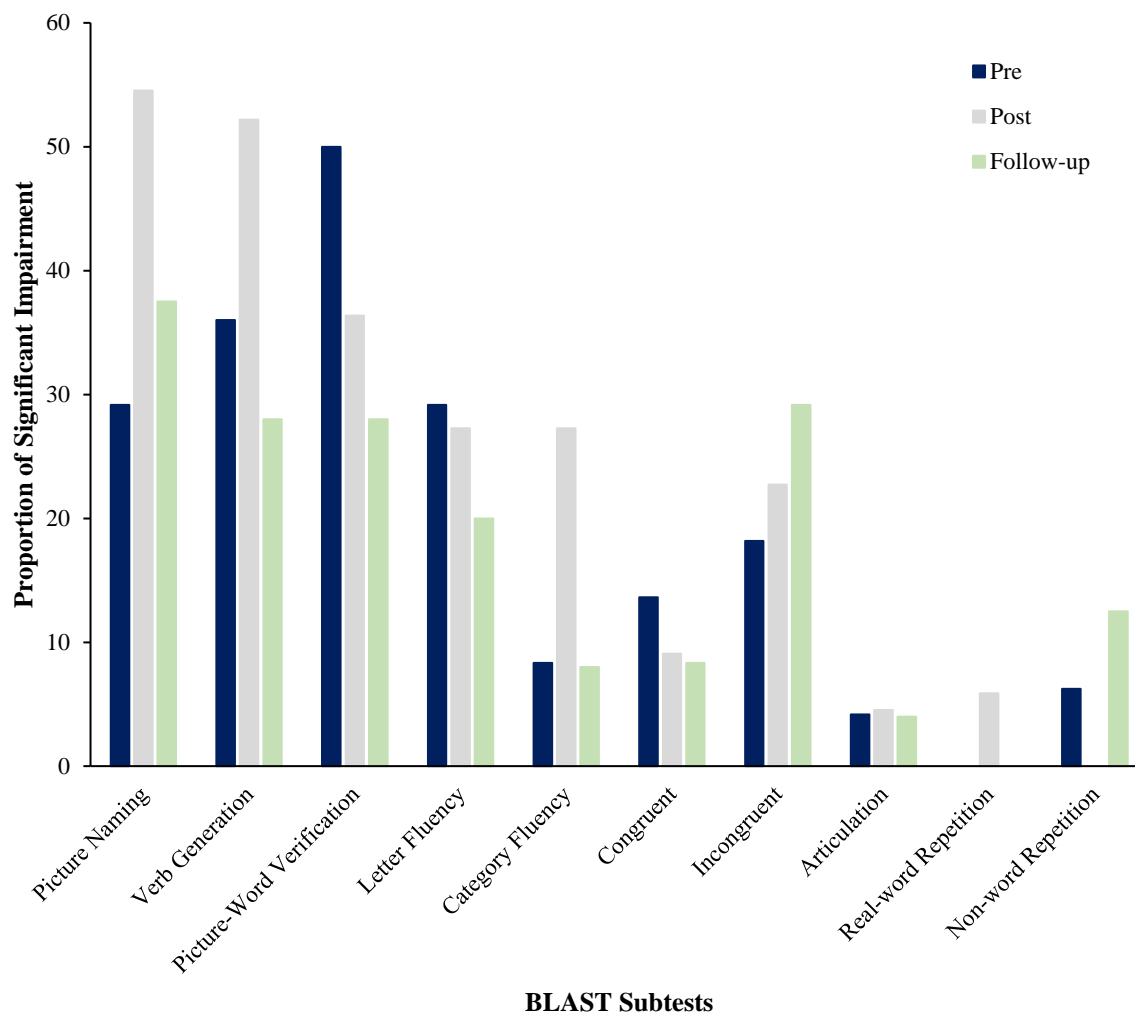
Figure 2.8. The proportion of patients significantly impaired on the tasks contained in the BLAST across the three surgical phases using Crawford & Howell's (1998) analysis. Higher bars represent higher proportion of significant impairment, whilst lower bars represent less significant impairment.

Preoperatively, a surprisingly high 95% of patients scored significantly below their respective control groups on at least one task on the BLAST. Postoperatively, the corresponding figure was 92%, and at long-term follow-up, it was a slightly better 78%. Looking across all three testing phases, *Picture Naming* appeared to be the most commonly failed task, followed by *Picture-Word Verification*, *Verb Generation*, and *Letter Fluency*. In

general, tasks that were sensitive at detecting impairment during the acute surgical phase continued to be sensitive during the follow-up period.

As shown in Figure 2.8., between the acute preoperative and postoperative phases, the number of patients scoring below normal declined on some tasks. This was the case for *Picture-Word Verification*, *Letter Fluency*, *Non-Word Repetition*, and the *Congruent* condition of the Stroop task. In contrast, there was actually an increase in the number of patients scoring below normal on the following tasks: *Verb Generation*, *Picture Naming*, *Category Fluency*, the *Incongruent* condition of the Stroop task, *Real-word Repetition*, and *Articulation*.

Between the acute surgical stage (preoperative and postoperative) and long-term follow-up, some patients' scores improved to normal levels. Specifically, there were fewer patients who scored below normal on the following subtests: *Picture-Word Verification*, *Verb Generation*, *Letter Fluency*, *Category Fluency*, and the *Congruent* condition of the Stroop task. However, not all tasks elicited improved performance at long-term follow-up. Relative to preoperative and postoperative testing, there were greater rates of significant impairment on *Non-Word Repetition*, and the *Incongruent* condition of the Stroop task.

Within-subject change. We next explored the nature of change across the surgical phases. To do this, each individual's accuracy score on each task was standardised based on the mean and standard deviation of the relevant control group, using the following formula: T (= (*individual* score – control group *Mean*)/ control group *Standard* Deviation)x10+50)[10] (see Miller & Rohling, 2001). This method ensures that the patient's score is expressed relative to what would be expected given their age[11]. To investigate the nature of change within individuals, we recorded any changes in T scores of at least 20 points from preoperative testing to follow-up; this difference is the equivalent of two standard deviations, and is considered a conservative reflection of change. Impaired preoperative performance was defined as any score that fell below 30 T scores, and is considered two standard deviations below the mean – this criterion resulted in very similar rates of individual impairment to Crawford and Howell's (1998) modified t test. An increase of at least 20 points (T scores) across time periods was defined as a *significant improvement*, and a decrease of at least 20

---

[11] The advantage of this process is it provides a standardised measure that allows for a normal distribution and enables for the comparison of scores that are from different distributions. T scores can also better control for age-related effects by standardising scores relative to age-matched controls.

points (T scores) was defined as *significant deterioration.* All other outcomes (e.g. < +/- 20 T scores) were defined as *no change*.

For this section, we chose not to examine immediate post-operative performance, as scores at the timepoint may be heavily influence by extraneous variables such as postoperative swelling, fatigue, and the post-operative effects of anaesthesia and other medication (see Heimans & Reijneveld, 2012). Also, we opted not to include our repetition subtests when examining within-subject change, as a high number of patients did not complete preoperative testing on these tasks due to time constraints and fatigue.

Figure 2.9. displays the relevant results. As is evident from the Figure, a number of patients exhibited newly acquired significant impairments during the follow-up phase, which was most evident on *Picture Naming, the Incongruent* condition of the Stroop task*, and Verb Generation.*

As seen in Figure 2.9., there were instances where patients' scores improved substantially at follow-up, according to the 20-point change criterion. The skills that yielded the highest incidence of improvement were *Picture-Word Verification* (eight cases), and *Verb Generation* (five cases). In these cases, the relevant scores were impaired preoperatively. On the other hand, however, there was a surprisingly high number of instances where scores substantially deteriorated, despite unimpaired preoperative performance. The skills that yielded the highest rates of deterioration were *Picture Naming* (four cases), the *Incongruent condition* of the Stroop task (four cases), and *Verb Generation* (three cases). In all of the cases of deterioration, the individual's score was within the normal range at preoperative testing.

In terms of specific factors that may give rise to this long-term decline in performance, it is worth noting that no specific patient, nor tumour localisation appeared to be driving these effects. Importantly however, of the 13 patients who showed a substantial long-term decline on at least one task, despite their unimpaired preoperative performance, 67% had meningioma, whilst 15% had low-grade glioma. This observation suggests that long-term outcomes for those with low-grade tumours may not be as favourable as previously thought.
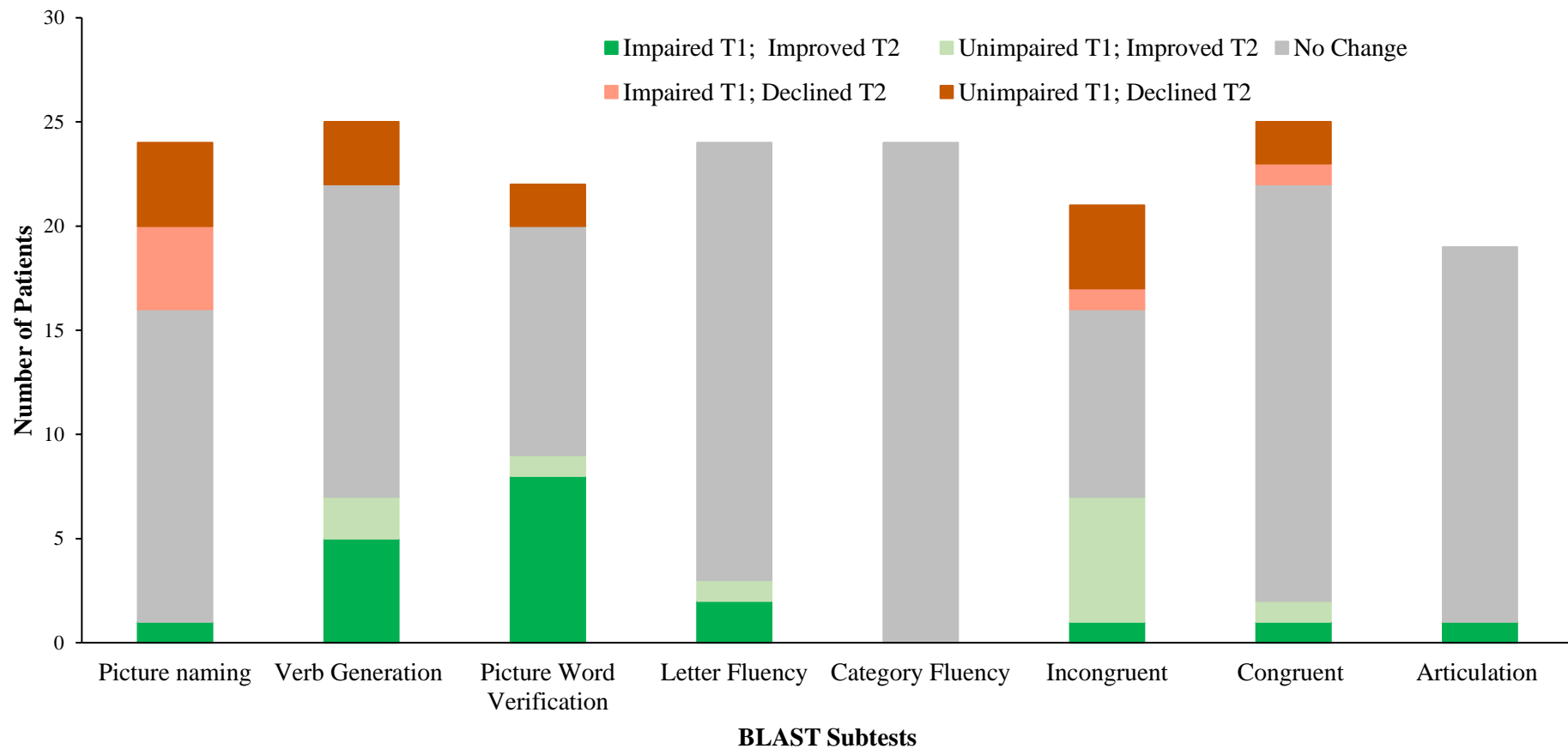
Figure 2.9. Total numbers of participants whose skill scores substantially improved (>20 point increase in T score), declined (>20 point decrease in T scores), or elicited no significant change (< +/- 20 T scores) from preoperative to follow-up testing phases. Higher bars reflect a greater proportion of impairment (defined as a T score less than 30), whilst lower bars reflect a lower proportion of impairment. T1 = time one (pre-operative); T2 = time two (follow-up) – for example, Impaired T1; Improved T2 indicates that patients were impaired during the acute pre-operative phase (<30 T score), however their scores on the respective subtest improved substantially (> 20 t scores) at long-term follow-up.

To explore whether preoperative scores predicted performance at long-term follow up, we performed a linear regression using SPSS Software. The relevant preoperative scores were entered as the predictor variable and the corresponding follow-up scores as the dependent variable. As above, we did not examine immediate post-operative scores due to the various extraneous factors that may impact on performance. At the individual task level, preoperative scores on the *Category Fluency* task reliably predicted scores at long-term follow-up ($F(1,22)= 30.785$, $p< .000$, $R^2= .583$). The corresponding analysis for the *Picture Naming* task just failed to reach statistical significance ($F(1, 22)= 4.091$, $p= .055$, $R^2 = .157$,). No other preoperative scores were significant predictors of long-term follow-up performance.

Effects of tumour localisation. Figure 2.10. displays the percentage of patients who exhibited significant impairments on the most sensitive BLAST subtests as a function of the primary localisation of the tumour: *right frontal (6)*, *right posterior (4)*, *left frontal (4)*, *left parietal (6), left temporal (4)*[12] (Classification of these localisations are outlined in Chapter 2, section 2.1.1). As above, significant impairment was based on Crawford and Howell's (1998) modified t-test.

Pre-Operative

---
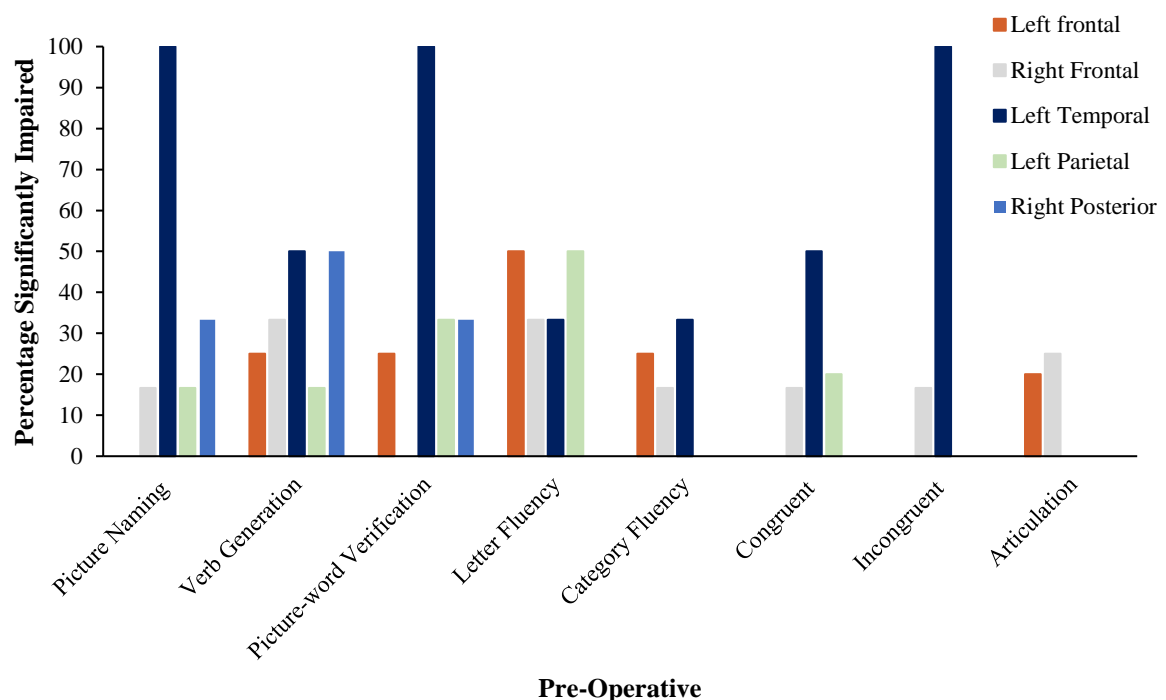
[12] One patient, AE, was excluded from this analysis due to the presence of multiple lesions
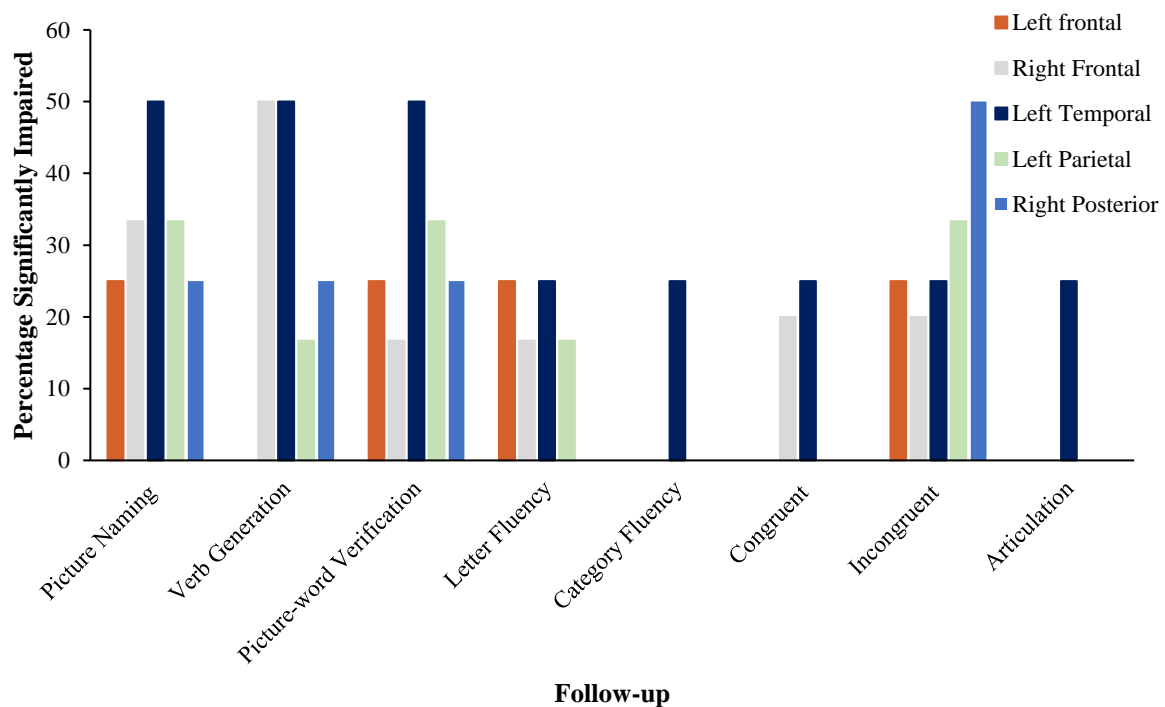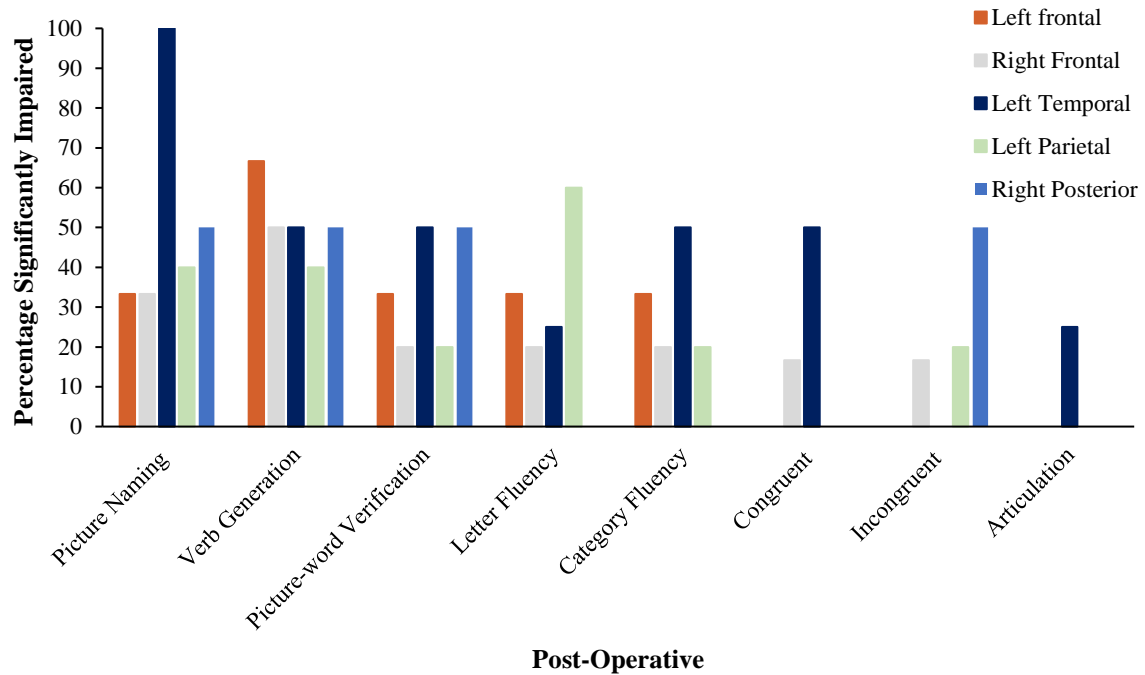
Figure 2.10. The percentage of patients with significantly impaired accuracy (according to Crawford & Howell's (1998) modified t-test) according to tumour localisation across the pre-operative, post-operative and follow-up phase. Higher bars reflect higher rates of significant impairment, whilst lower bars reflect lower rates of significant impairment.

To investigate whether significant differences exist between tumour localisation and accuracy performance on the BLAST subtests, we examined whether there were significant group differences with respect to the patient's average T scores for each subtest. For each subtask and for each testing phase, we first performed an omnibus one-way Analysis of Variance (ANOVA) using SPSS software, considering all groups simultaneously, and then if there was a significant overall group effect, we specifically compared scores for the group of interest with those for the other groups.

Preoperatively, there was a significant effect of tumour localisation on *Picture Naming* scores ($F(4,22)= 3.592$, $p= .025$). Given we did not make specific predictions with regards to the BLAST subtests, we then performed a Tukey's post-hoc analysis which revealed that left temporal patients ($M= 6.737$, $SD= 35.97$) scored significantly lower than left frontal patients ($M= 51.13$, $SD= 2.99$). A significant effect was also obtained for the *incongruent condition* of the Stroop task ($F(4,20)= 6.204$, $p= .003$), with a Tukey's post-hoc analysis revealing that those with left temporal tumours ($M= 9.14$, $SD= 10.32$) performed significantly more poorly than those with left frontal tumours ($M= 55.03$, $SD= 1.12$). Such discrepancy may be due to left temporal patients having difficulty retrieving the word form from their lexicon, which is discussed further in later Chapters (see Chapters 4 and 5). No other significant effects were obtained.

During the immediate post-operative and long-term follow-up phases, there was no significant group effects of tumour localisation group on any of the task scores. Therefore, in no instance did we proceed with more specific hypothesis testing.

Importantly, these statistical results should be treated with caution because the numbers of patients in each localisation group were small, and the number of statistical tests performed was high (24 separate analyses of variance in total). Given the high number of tests performed, it may become appropriate to apply a Bonferroni correction to the *p* threshold; the resultant *p* value would then be .002. According to this criterion, no analyses would have reached the threshold for statistical significance.

Investigation of Core Cognitive Language Skills

Turning now to the main objective of this study – which was to examine each of the eight cognitively-defined core skill measures across the different testing phases. These scores were calculated from each patient's performance on the key tasks/subtasks of the BLAST. Table 2.9. outlines each cognitive skill and the key performance measures that contribute to their operationalisation. Consistent with the approach utilised by Faulkner and colleagues (2017), when possible, each of the contributing performance measures were first expressed as a z score, using the mean and standard deviation from that patient group as the reference values. The primary purpose of converting to z scores in this way was to ensure that all aggregated measures used comparable units and had a similar spread of distribution. For these purposes, it is preferable to use the patients themselves as the reference sample, rather than the relevant age-matched control sample, because the latter often have a very limited range (for example, some measures were subject to ceiling effects). Accordingly, this process ensured that each of the measures were weighted equally, regardless of the measurement scale (see Faulkner et al., 2017). Following this process, scores were then combined into the appropriate formulae (see Table 2.9) to derive a total score for that skill. The fourth column in Table 2.9. presents the relevant formula for each core skill measure, which were combined to derive a total z score.

Given the confounds that are present with the acute post-operative phase, our examination of the core skills includes only the pre-operative and follow-up periods. This method is identical to the one used by Faulkner and colleagues (2017), who exclusively examined immediate pre-operative performance.

Table 2.9.

Summary of the cognitive skills, and their associated measures (see also Faulkner et al., 2017)

Cognitive Skill	BLAST Profile	Key Performance measures	Formula
Accessing Semantic Knowledge	Semantic confusions on picture-word verification Production of semantic errors in picture naming Poorer category fluency relative to letter fluency	Percent of semantic errors on picture naming plus percent of semantic confusion errors in picture-word verification, expressed as a standardised score (individual score – patient group mean/standard deviation of patient group) (c) Category fluency score minus letter fluency score, expressed as a standardised score (d)	Mean (c,d)
Lexical Selection	Frequency effect on picture naming Disproportionately high production of semantic paraphasias and omissions in picture naming The above not accounted for by a general semantic deficit (as indicated by high rates of semantic confusions in auditory comprehension, and disproportionately low category fluency scores)	Slope of line that expresses the relationship between frequency and accuracy in picture naming, converted into a standardised score (e) Percent of omissions in picture naming, expressed as a standardised score (f) Percent of semantic confusions in picture-word verification, expressed as a standardised score, plus difference between category and letter fluency score, expressed as a standardised score (g)	Mean (e, f), minus any variance shared with g
Phonological encoding	Strong length effects in picture naming	Slope of line that expresses the relationship between length and accuracy in picture naming, expressed as a standardised score (h)	Mean (h, i) minus any variance shared with (j)

	Production of phonological errors in picture naming	(Percent phonological errors in picture naming + percent errors in real word repetition + percent of errors in non-word repetition), expressed as standardised score (i)	
	Errors in single word and non-word repetition	Total score in articulatory agility expressed as standardised score (j)	
	The above not accounted for by poor articulatory agility alone		
Auditory Word Recognition	Impaired single word repetition	(Percent errors in word repetition +Percent phonological confusion errors in picture- word verification), expressed as a standardised score (a)	Mean (a, b)
	Phonological confusions in picture-word verification		
		Slope of line that expresses the relationship between word length and accuracy, converted into a standardised score (b)	
	Reverse length effect in real-word repetition		
Goal-Driven Response Selection	On the Stroop task, disproportionately high error rates and/or slow response times on incongruent items	Increase in RT from congruent to incongruent items, expressed as standardised score (l)	Mean (l, m, n, o) minus any variance shared with (p)
	On the verb generation task, disproportionately high error rates on high selection items	Difference in percentage of errors on incongruent and congruent items, expressed as a standardised score (m)	
	Poor letter fluency scores	Difference in percentage of errors on low and high selection items, expressed as a standardised score (n)	
		Letter fluency score, expressed as a standardised score (o)	
	The above not accounted for by a general naming deficit (as indicated by poor naming of high frequency objects	Percent errors in picture naming task (high frequency	

		items only), expressed as a standardised score (p)	
Verb Retrieval	Disproportionately impaired on high selection items from verb generation task The above not accounted for by a general naming deficit (poor naming of high frequency objects)	Percent errors in verb generation (high selection only) expressed as a standardised score (q). Percent errors in picture naming (high frequency items only), expressed as a standardised score (r)	q minus variance due to r
Verbal STM	Poor non-word repetition	Percent errors in nonword repetition, expressed as standardised score (k)	k
Articulatory-Motor Planning	Poor performance on BDAE articulatory agility task	Articulatory agility score, expressed as a standardised score (j)	j

Our first objective was to investigate the overall incidence of abnormally low scores on each of the core skill measures. Each patient's accuracy score (first standardised relative to the entire patient group) on each of the eight core skills was compared with that of the respective control group using Crawford and Howell's (1998) modified t-test. For the purpose of this analysis, we only assessed core skills if patients had completed sufficient tasks to formulate an overall score. Figure 2.11. shows the overall incidence of skill scores that were significantly below those of their relevant control group in each testing phase, based on these analyses.
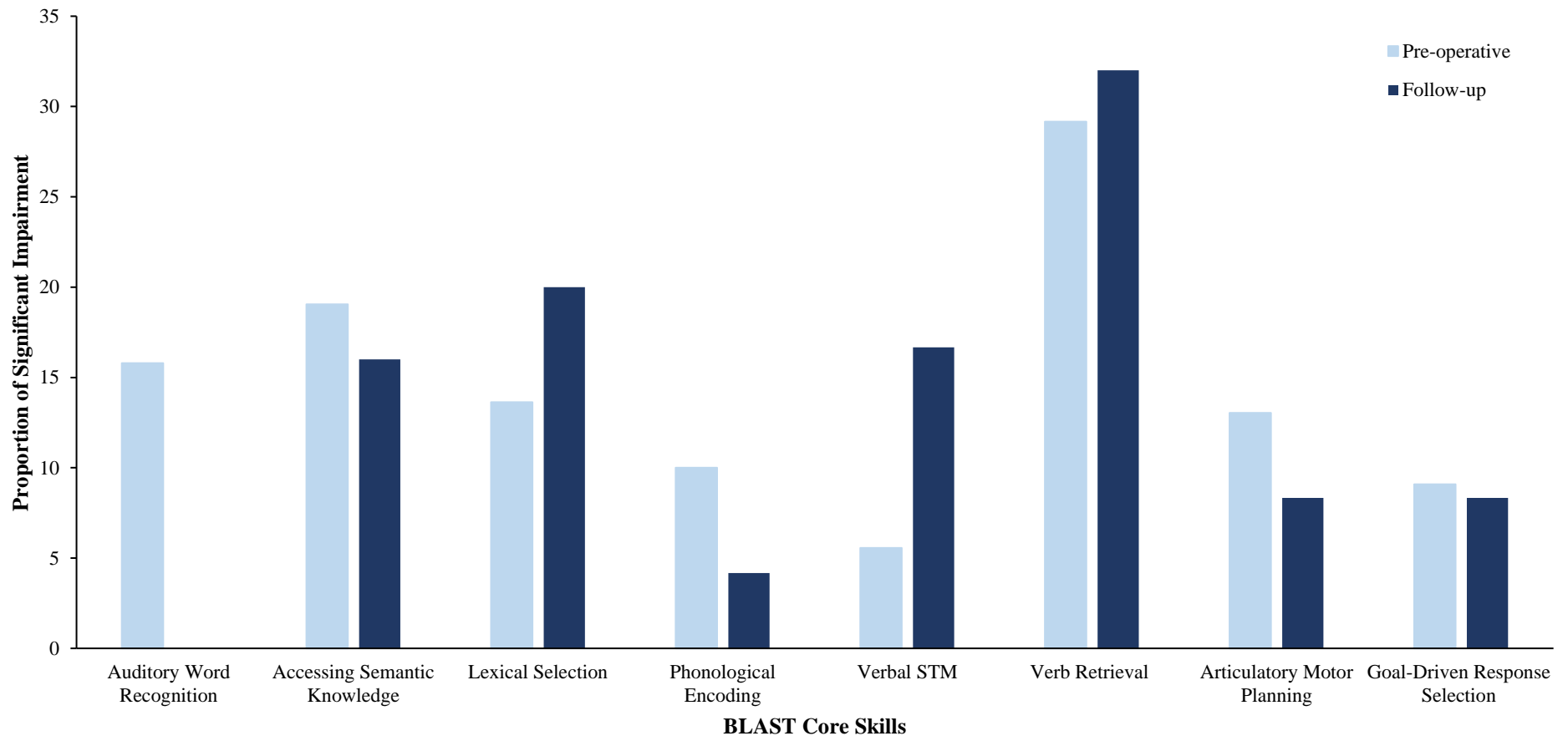
Figure 2.11. Proportion of patients with significant impairment on each core cognitive skill across the pre-operative and follow-up phases according to Crawford and Howell's (1998) modified t test. The x axis reflects the BLAST core skills, whilst the y axis reflects the proportion of significant impairments. Higher bars reflect a greater proportion of patients with significant impairments, whilst lower bars represent a lower proportion of patients with significant impairments.

Pre-operatively, the incidence of significant impairment was surprisingly high, with 71% of patients scoring significantly below their respective control group on at least one core language skill. At follow-up, the incidence of significant impairments was lower, albeit still surprisingly high, with 49% of patients scoring significantly below their respective control group on at least one skill measure. As is evident from Figure 2.11, during the preoperative phase, *verb retrieval* was the most commonly impaired core skill (29% of patients impaired), followed by *accessing semantic knowledge* (19%), whilst *auditory word recognition* (16%) was also particularly sensitive during this phase.

A similar trend was evident at follow-up, with *verb retrieval* serving as the most commonly impaired core skill (32% of patients were impaired relative to their appropriate control group). However, there was also a high incidence of impairment on *lexical selection* (20%) and *verbal short-term memory* (17%).

Within-subject change.

We next explored the nature of change across the preoperative and follow-up phases. As above, a substantial change was defined as any change in T scores of at least 20 points, irrespective of whether the preoperative score was significantly impaired or not. The value of 20 was selected as it is equivalent to two standard deviations, which is a conservative reflection of change. Appendix D presents the core skill scores for each individual patient, and also highlights which skill scores substantially improved or declined according to the 20-point criterion.

Figure 2.12. shows the nature of performance change across the two phases. As is evident from the Figure, a number of patients exhibited newly acquired significant impairment during the follow-up phase, which was most evident on *verb retrieval* and *lexical selection*. Specific factors that may account for this finding will be discussed in subsequent sections.

Also seen in Figure 2.12, there were some instances where patients' scores improved substantially at follow-up, according to the 20-point change criterion. The skills that yielded the highest incidence of improvement were *verb retrieval* (three cases) and *auditory word recognition* (three cases). In all of these cases, the relevant scores were impaired preoperatively (<30 T scores). Moreover, *articulatory motor planning* (two cases), and *goal-driven language selection* (two cases) showed higher rates of improvement compared to the other skill measures. On the other hand, however, there were even more instances where

scores substantially deteriorated at follow-up. The skills that yielded the highest rates of deterioration were *verb retrieval* (five cases) and *lexical selection* (four cases). In all of the cases, the individual's score was within the normal range at preoperative testing. As above, such findings suggest that long-term language performance may not be as optimal as previously thought.
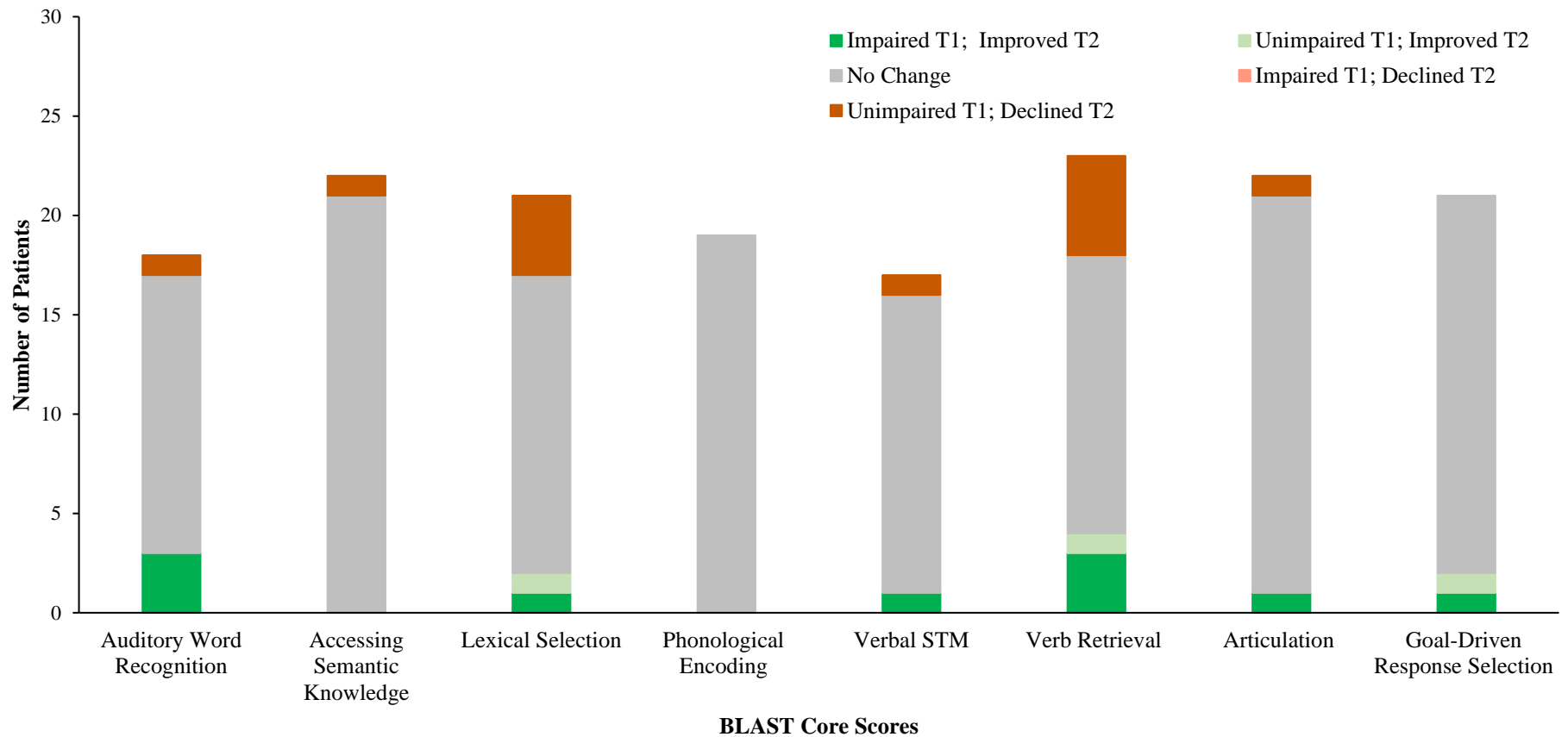
Figure 2.12. Total numbers of participants whose skill scores substantially improved (>20 point increase in T score), declined (>20 point decrease in T scores), or elicited no significant change (< +/- 20 T scores) from preoperative to follow-up testing phases. The x axis represents the BLAST core skills, whilst the y axis represents the number of patients. Higher bars reflect a greater proportion of impairment (defined as a T score less than 30). T1 = time one (pre-operative); T2 = time two (follow-up).

In terms of specific factors that may give rise to this long-term decline in performance, it is worth noting that only 33% of patients who experienced a substantial decline on at least one skill measure presented with frontal tumours, whilst 44% presented with temporal tumours. Importantly, no one patient, tumour type or location appeared to be driving these significant deteriorations across the core skills.

To explore the nature of change more statistically, we also compared preoperative and follow-up scores for the entire patient group for each specific skill score. Using SPSS Software, we performed a paired samples t test, with preoperative scores serving as the independent variable, and follow-up test scores serving as the dependant variable. A significant effect of testing phases was found for *auditory word recognition* ($t(18) = 3.090$, $p= .006$) and *goal-driven response selection* ($t(20) = 3.140$, $p= .000$). In these instances, scores were significantly higher at follow-up. A significant effect of testing phases was also observed for *lexical selection* ($t(21) = 2.276$, $p= .033$), however in this instance, scores were actually significantly lower during the follow-up phase compared to pre-operative testing. This finding is particularly surprising as it suggests that scores generally *declined* for lexical selection between the acute phase and long-term follow-up. The relevant means and standard deviations are presented in Appendix D.

Given our expectation that higher-grade tumours would result in greater follow-up decline, we next assessed the extent to which tumour type may account for deterioration in performance scores across the pre-operative to follow-up period. We first started with a very general approach by categorising patients according to high grade (n= 5) and low grade (n = 20) lesions (for further details of this, see section 2.1.1). Table 2.10 displays the relevant group means and standard deviations for each skill score. Using an independent samples t test, we first explored whether there were any significant differences between the two groups for each of our skills scores. Looking first at preoperative scores, a significant effect was obtained for *lexical selection* ($t(20) = 3.051$, $p= .006$), with low-grade tumour patients (*M*= 56.19, *SD*= 12.74) scoring significantly higher than those with high-grade lesions (*M*= 28.9, *SD*= 28.8). No other significant effects were obtained. Turning now to follow-up scores, a significant effect was also obtained for *lexical selection* ($t(23) = 2.195$, $p= .039$). Again, those with low-grade tumours (*M*= 47.30 *SD*= 12.37) scored significantly higher on this skill measure than those with high-grade lesions (*M*= 31.84, *SD* = 16.09).

Table 2.10.

Means and standard deviations (in parentheses) for low and high-grade tumour types for each skill score, across the preoperative and follow-up phases.

Core Skills	Low (Pre-operative)	Low (Follow-up)	High (Pre-operative)	High (Follow-up)
Auditory Word Recognition	44.75 (*12.57*)	53.12 (*5.92*)	39.52 (*19.03*)	54.17 (*4.06*)
Accessing Semantic Knowledge	44.29 (*12.06*)	46.46 (*11.79*)	38.67 (*10.43*)	42.11 (*12.87*)
Lexical Selection	56.23 (*13.13*)	47.14 (*12.66*)	34.22 (*27.64*)	35.61 (*16.28*)
Phonological Encoding	49.88 (*13.24*)	47.06 (*14.15*)	50.55 (*12.00*)	46.94 (*7.63*)
Verbal Short-term Memory	46.96 (*16.12*)	44.89 (*15.78*)	55.68 (*4.50*)	54.90 (*4.93*)
Verb Retrieval	42.32 (*21.98*)	39.92 (*21.61*)	44.27 (*16.23*)	39.74 (*36.32*)
Articulatory Motor Planning	44.06 (*12.48*)	44.85 (*11.16*)	40.72 (*7.70*)	42.49 (*6.85*)
Goal-driven Response Selection	43.59 (*7.63*)	50.79 (*8.38*)	40.52 (*14.26*)	50.81 (*14.04*)

We next explored whether each patient's mean difference was statistically significant between the two testing phases. Considering low-grade tumours first, a paired samples t test revealed a significant effect of testing phases for *auditory word recognition* ($t(16) = 2.704$, $p= .016$) and *goal-driven response selection ($t(17) = 3.049$, $p= .007$).* In both instances, patients with low-grade tumours scored significantly higher at follow-up compared to their respective preoperative scores. A significant effect of testing phases was also revealed for *lexical selection* ($t(17) = 3.910$, $p= .001$). However, in this case, patients with low-grade tumours actually scored significantly *worse* at follow-up compared to their preoperative scores.

Turning now to high-grade tumours, a paired samples t test revealed no significant effect of testing phases. These non-significant effects suggest that patients did not significantly differ on any core skill between the two periods. This may be attributed to the very small sample size in the high-grade tumour group (n=5).

Exploratory analyses were also performed to determine whether any significant differences existed between the 25 patients included in the current sample and those who passed away prior to follow-up testing. To do this, we performed an independent samples t test to statistically compare pre-operative scores on each of our core skills measures for our two groups. A significant effect was obtained for *goal-driven response selection* ($t(20) =$ 3.051, *p*= .009), with those who passed away scoring significantly more poorly than those who were able to participate at long-term follow-up. Importantly, this effect was not accounted for by tumour volume or location, nor patient age. Possible explanations for this finding are discussed in the subsequent section.

Our next objective was to investigate whether pre-operative scores served as a significant predictor of follow-up performance. To do this, we performed a linear regression using SPSS Software. The relevant preoperative scores were entered as the predictor variable and the corresponding follow-up scores as the dependent variable. Pre-operative scores on the following skills were significant predictors of follow-up performance on their respective scores: *accessing semantic knowledge* ($F(1, 21) = 11.076$, *p*= .003, $R^2 = .345$), *lexical selection* ($F(1, 20) = 17.937$, p= .003, $R^2 = .473$), *phonological encoding* ($F(1, 17) = 27.288$, *p*< .000, $R^2 = .616$), *verbal short-term memory* ($F(1,15) = 5.034$, *p*= .040, $R^2 = .251$). Whilst not statistically meaningful, *verb retrieval* was approaching significance ($F(1, 22) = 4.238$, *p*= .052, $R^2 = .162$). Figure 2.13 displays the relevant scatterplots for each of these significant results.
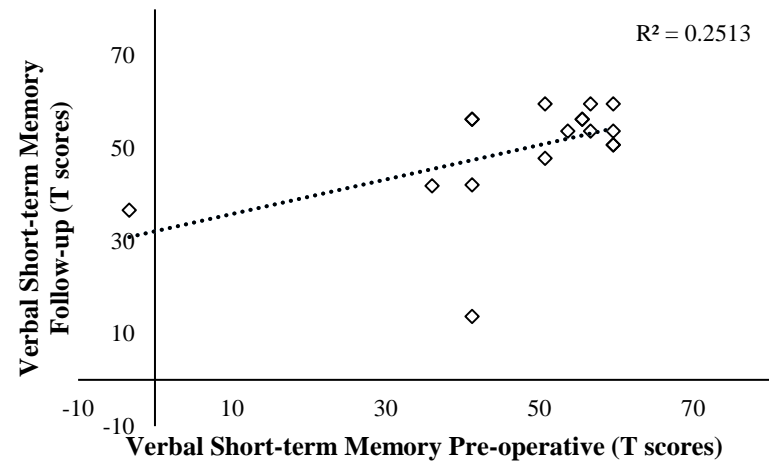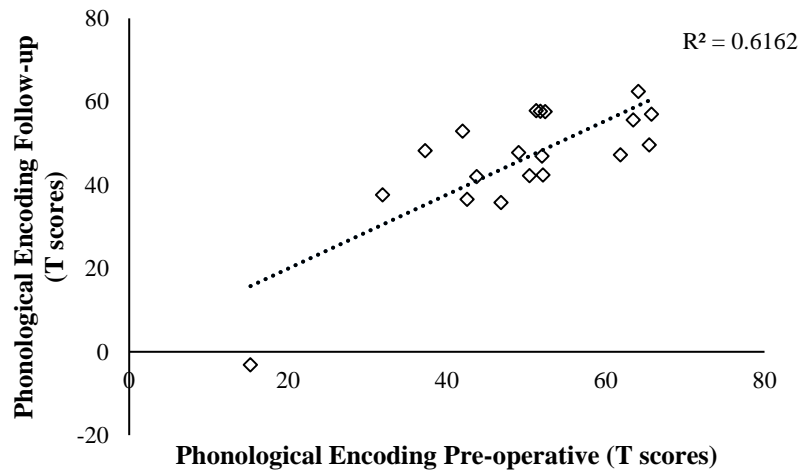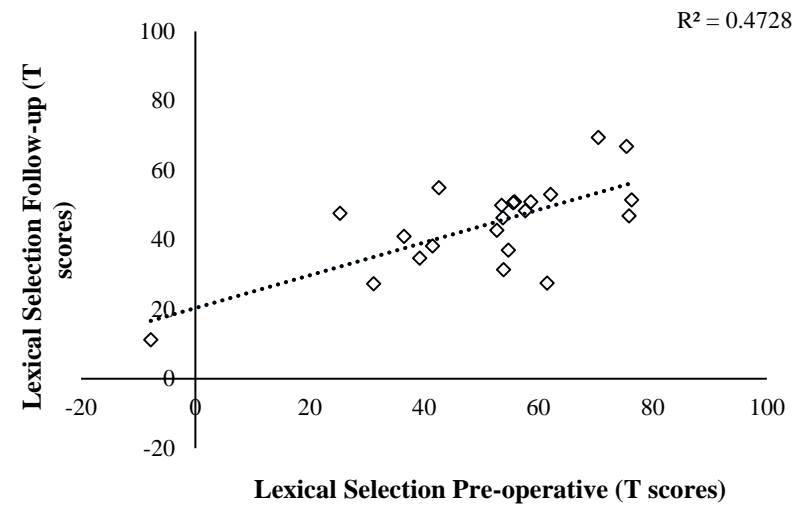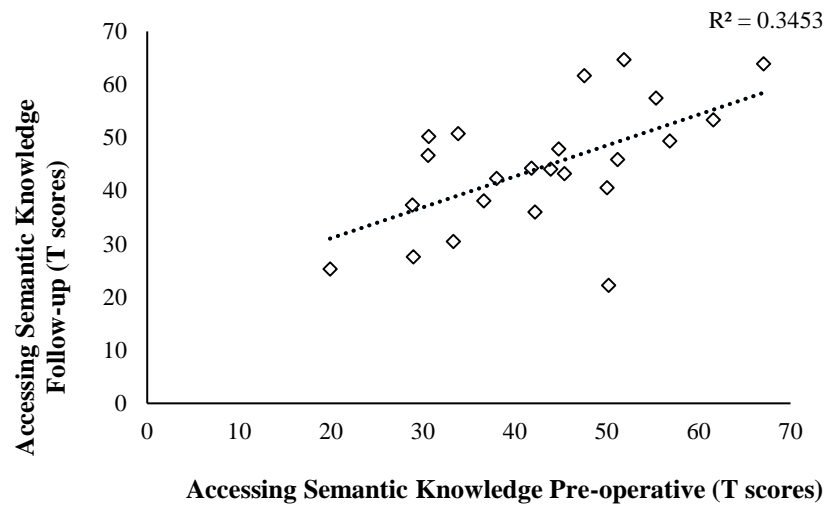
Figure 2.13. Individual scatter plots for significant core skill predictors across the pre-operative and follow-up periods (in T scores).

Factors affecting follow-up performance

Turning now to specific factors that may affect follow-up performance. We investigated whether the following were significant predictors of performance change across the pre-operative to follow-up period: i) *tumour type and grade;* ii) *tumour volume (cm3);* and iii) *chronological age at the time of follow-up*. As with the subtests, performance change was defined as the absolute change in T scores for each core skill from the pre-operative to follow-up period, irrespective of whether this difference was significant.

Preoperative performance. We explored whether an individual's pre-operative scores were significantly predictive of the extent of follow-up change – that is, can preoperative performance predict how much improvement patients are likely to show at long-term follow-up. To explore this, we performed a linear regression using SPSS Software, with preoperative scores entered as the independent variable, whilst absolute change scores (T scores, irrespective of whether the change was significant) served as the dependant variable. To determine the directionality of any effects, we also performed a Pearson's correlation.

Linear regression revealed a significant effect for the following skills scores: *auditory word recognition* ($t(1, 18) = 81.86$, $p< .000$ $R^2 = .83$), *accessing semantic knowledge* ($t(1, 22) = 5.71$, $p = .026 = R^2 = .21$), *lexical selection* ($t(1, 21) = 22.586$, $p< .000$ $R^2 = .53$), *verbal short-term* memory ($t(1, 15) = 13.196$, $p= .003 = R^2 = .48$), *verb retrieval* ($t(1, 23) = 4.973$, $p= .036$ $R^2 = .18$), *articulatory motor-planning* ($t(1, 21)= 22.404$, $p <.000$ $R^2 = .53$), and *goal-driven response selection* ($t(1, 20) = 22.62$, $p< .000$. $R^2 = .54$). In all instances, the directionality of these effects was negative, which suggests that lower preoperative scores were significantly predictive of greater follow-up improvement – in other words, the lower the pre-operative score, the more likely it was to show greater improvement at long-term follow-up. The only core skill that did not show a significant effect was phonological encoding, wherein no patients showed any substantial change at long-term follow-up.

Tumour volume. We predicted that higher tumour volume (cm³) would be associated with greater improvement on the various core skills. To explore this, we performed a linear regression on SPSS Software, with tumour volume (cm³) entered as the continuous independent variable, and absolute change scores (T scores, irrespective of whether the change was significant) entered as the dependant variable. Overall, tumour volume (cm³) was a significant predictor of change across the pre-operative to follow-up period for: *accessing semantic knowledge* ($F(1,20) = 7.671$, $p= .012$, $R^2 = .277$); and *articulatory motor planning* ($F(1,18) = 7.895$, $p= .012$, $R^2 = .305$). In both instances, higher tumour volume was actually

associated with *less* overall improvement across the preoperative to follow-up period. This finding was inconsistent with our original prediction, and is discussed further in the subsequent discussion section.

Tumour type. We predicted that lower-grade tumours would be associated with greater overall improvement than higher-grade lesions, such as glioblastoma and high-grade glioma, where we may even observe a deterioration in performance across phases. We performed a linear regression on SPSS Software to examine the effect of tumour type (meningioma, astrocytoma, low-grade glioma, high-grade glioma, and glioblastoma; coded as Grade I-V, which served as the independent variable) on overall improvement (in T scores, which served as the dependant variable) on each of the core skills scores. However, inconsistent with our original predictions, there was no significant effects of tumour type for any of the core skill measures.

Chronological Age. We predicted that age would be a significant predictor of overall improvement rates, with younger individuals exhibiting greater improvement, and/or less follow-up decline than older ones. To examine this, we performed a linear regression where each patient's age was entered as the continuous independent variable, whilst absolute change across the preoperative and follow-up (in T scores) was entered as the dependant variable. A significant, positive effect was observed for the skill measure of *auditory word recognition* ($F(1, 18) = 8.44$, $p= .010$, $R^2 = .332$). In this instance, greater age was actually associated with greater *improvement* between the preoperative and follow-up testing. This result was inconsistent with our original prediction, and may be due to younger patients showing a ceiling effect during the preoperative phase, which limited their potential to show substantial improvement at follow-up. Despite our original predictions, no other significant age effects were observed.

Effects of tumour localisation

Our next objective was to investigate our anatomical hypotheses with respect to the core skills measures. For these analyses, we examined preoperative scores and follow-up scores separately. Figures 2.14 and 2.15 display the incidence of significantly impaired scores (based on Crawford and Howell's (1998) modified t test), as a function of timepoint, for each of the following groups based on more defined lesion localisations: *left frontal* (4), *left temporal* (4), *left parietal* (6), *right frontal* (6), and *right posterior* (4). Figures 2.16 and 2.17

display the corresponding mean T scores for each core skill for each tumour group across the two surgical phases.
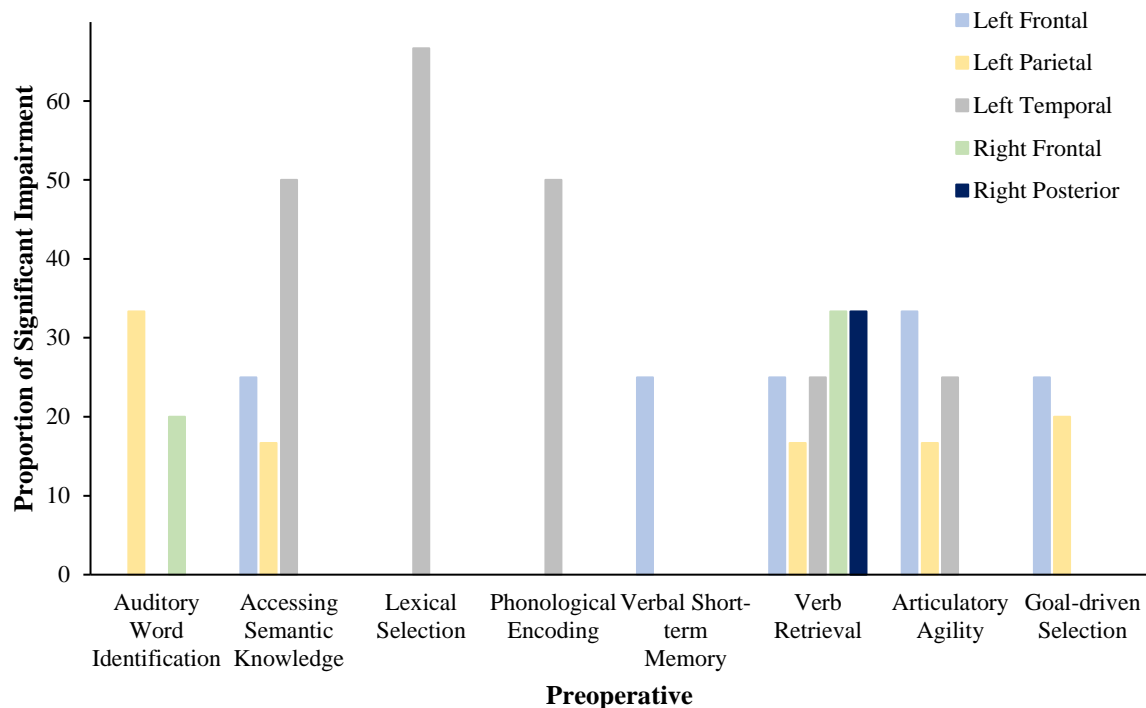
Figure 2.14. Proportion of significant impairment during the acute preoperative phase according to Crawford and Howell's (1998) modified t test. The x axis represents the pre-operative BLAST core skills, whilst the y axis represents the proportion of significant impairment. Higher bars reflect a greater proportion of significant impairment and lower bars reflect a lower proportion of significant impairment.
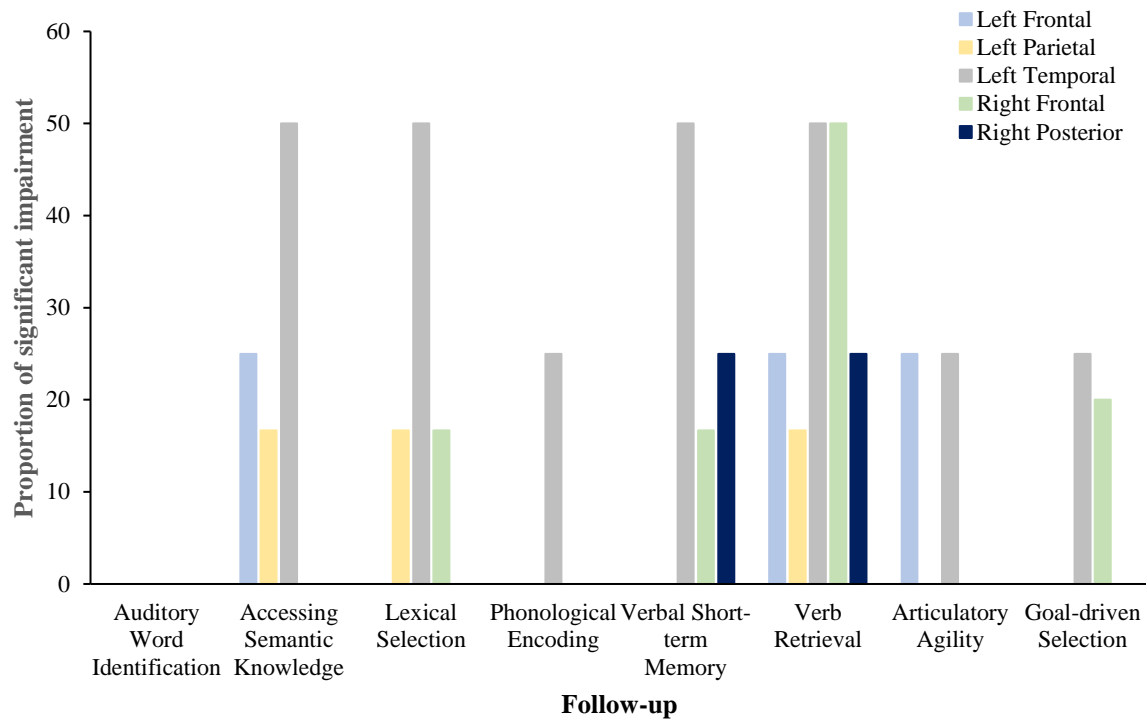
Figure 2.15. Proportion of significant impairment during the long-term follow-up phase according to Crawford and Howell's (1998) modified t test. The x axis represents the follow-up BLAST core skills, whilst the y axis represents the proportion of significant impairment. Higher bars reflect a greater proportion of significant impairment and lower bars reflect a lower proportion of significant impairment.
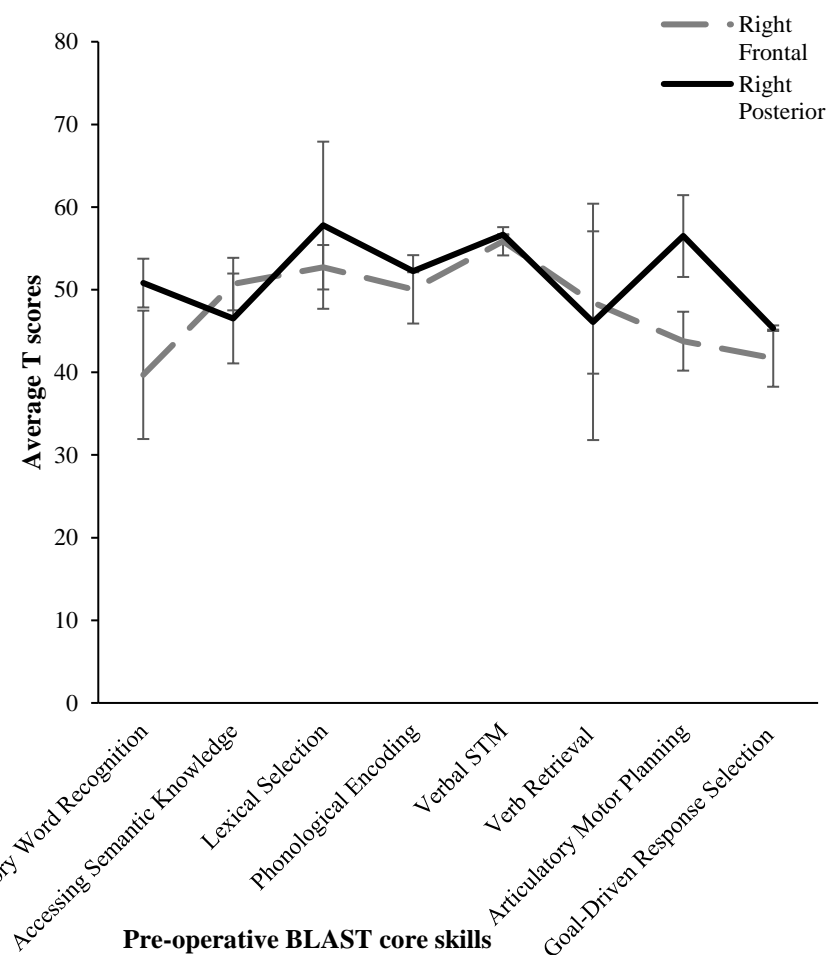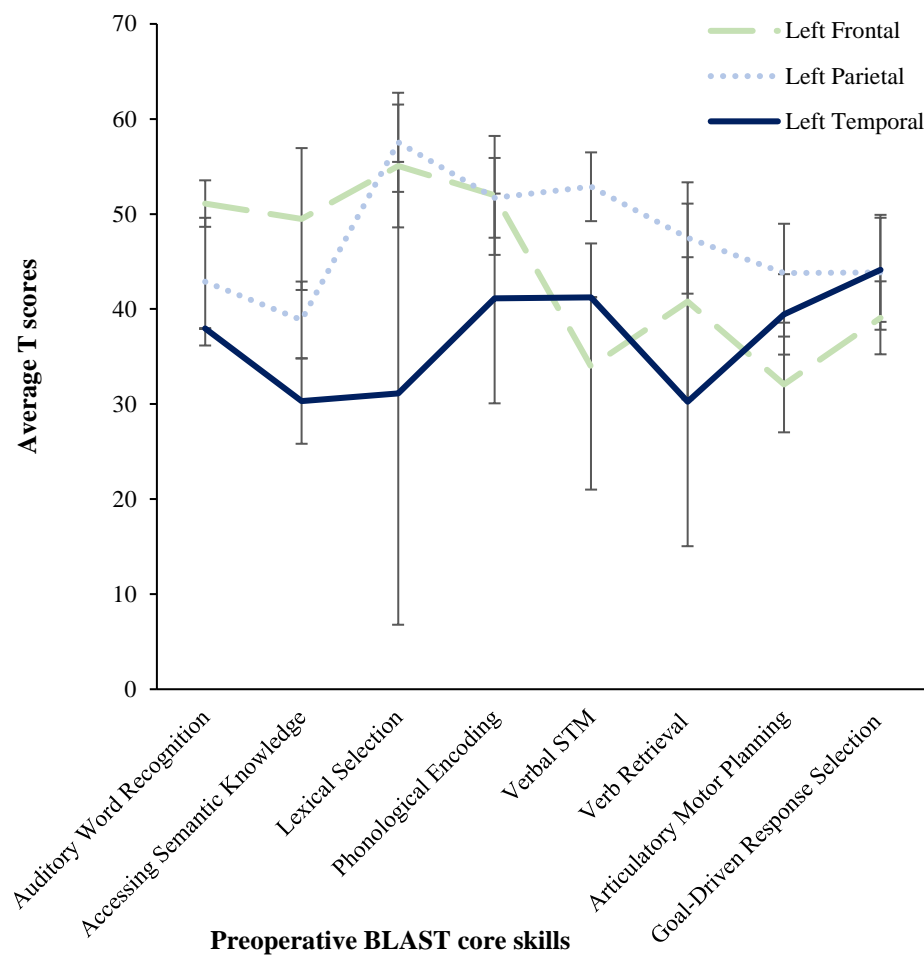
Figure 2.16. Average pre-operative T scores according to more specific tumour localisations. (N.B. error bars reflect standard error of the mean). Higher lines represent higher T scores, and better performance, whilst lower lines reflect lower T scores, and worse performance.
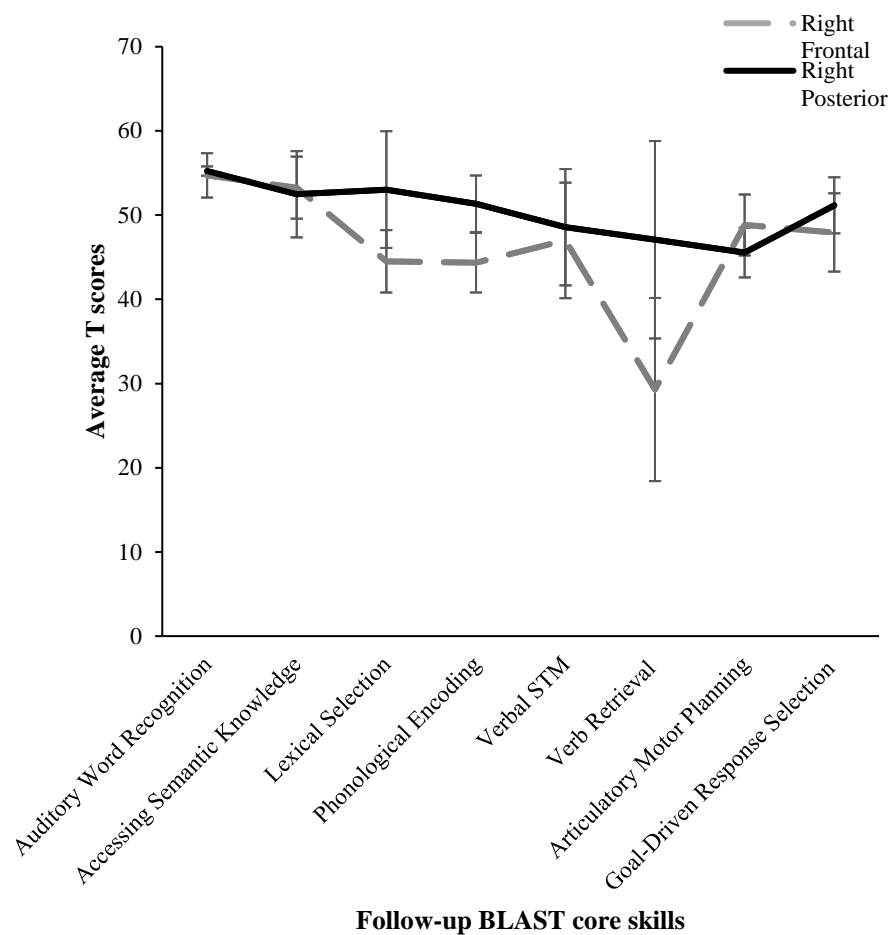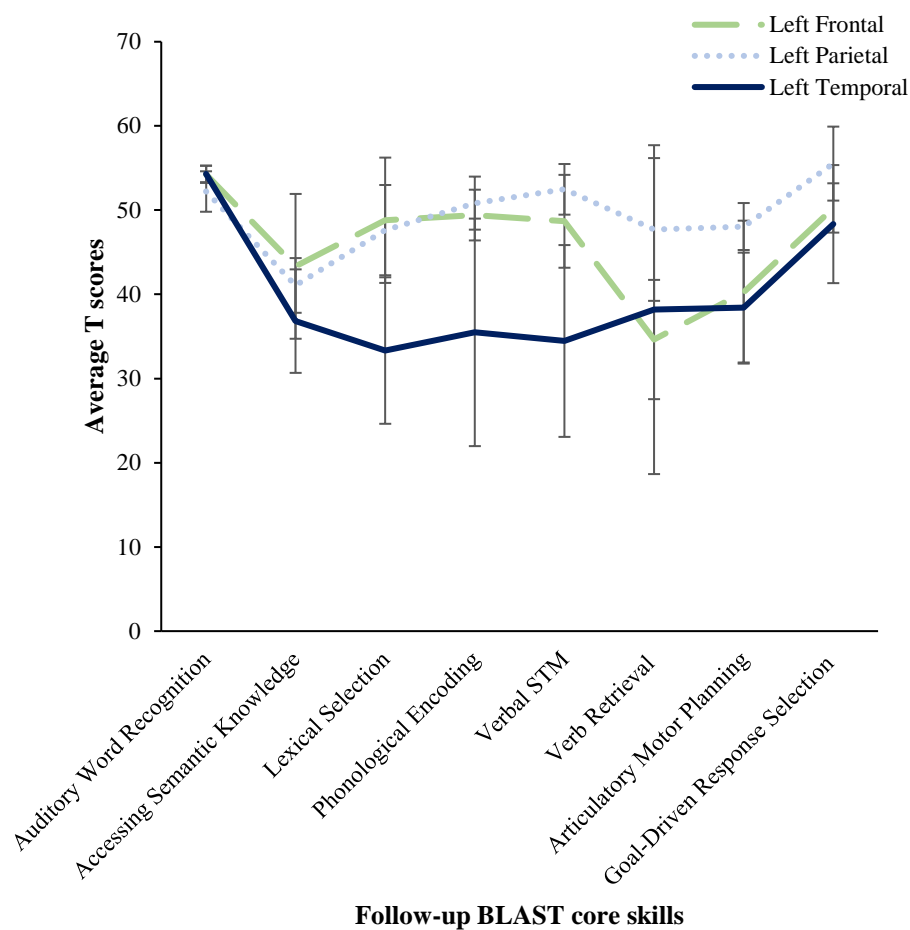
Figure 2.17 Average follow-up T scores according to more specific tumour localisations. (N.B. error bars reflect standard error of the mean).

Higher lines represent higher T scores, and better performance, whilst lower lines reflect lower T scores, and worse performance.

Again, considering the preoperative scores first, our first set of hypotheses was that left temporal patients would perform significantly more poorly than the remaining patients on: *accessing semantic knowledge, lexical selection, verbal short-term memory, phonological encoding,* and *auditory word recognition.* To test this hypothesis, we first performed an omnibus one-way ANOVA using SPSS Software, considering all groups simultaneously, and then if there was a significant overall group effect, we specifically compared scores for the group of interest with those for the other groups considered together. Overall, there were no significant effects of group, however *accessing semantic knowledge* just failed to meet statistical significance ($F(4,17)=$ .524, $p=$ .050).

Our second set of hypotheses was that left parietal patients would perform significantly more poorly than the remaining patients on measures of: *phonological encoding* and *verbal short-term* memory. However, using the same approach as above, a one-way ANOVA revealed no significant effects.

Our final set of hypotheses was that left frontal patients would score significantly more poorly than the remaining patients on: *goal-driven language selection, verb retrieval,* and *articulatory-motor planning*. Again, an omnibus one-way ANOVA revealed no significant effects between our tumour groups. Taken together, these non-significant findings are likely attributed to our small sample size and subsequent lack of statistical power to detect significant effects.

Turning now to the follow-up scores, we also performed an omnibus one-way Analysis of Variance for each of the core skills scores. However, none of these analyses yielded a significant group effect, although articulatory motor planning was approaching significance ($F(4,21)=$ 2.787, $p=$ .060). In no instance did we proceed with more specific hypothesis testing.

Importantly however, these statistical results should be treated with caution due to the small number of patients in each localisation group, and the high number of statistical tests performed (16 separate analyses of variance in total). As with the subtests, it may be more appropriate to apply a Bonferroni correction to the p threshold. Based on the number of analyses conducted, the resultant p value would be .003. According to this criterion, preoperative scores for accessing semantic knowledge would no longer reach statistical significance. Again, this non-significant effect is likely attributed to our small sample size and lack of statistical power.

Discussion

In this chapter, we examined tumour patients' performance on various language measures both preoperatively and at least three months after surgery. The main findings were as follows. First, in line with our predictions, and the results of previous studies, we observed improvements in performance between preoperative and follow-up testing phases on several language measures. At the subtest level, the tasks that yielded the highest rates of improvement were picture-word verification and verb generation. In these instances, the relevant scores were impaired preoperatively. On the BLAST core skills measures, those that yielded the highest rates of improvement were verb retrieval and auditory word recognition. Again, in these cases, the relevant scores were impaired preoperatively. However, not all measures showed improvement between the preoperative and long-term follow-up phases. Specifically, at the subtest level, there were greater rates of impairment on picture naming, non-word repetition, and the incongruent condition of the Stroop task when compared to preoperative scores. At the skills level, there were greater rates of impairment on lexical selection, verbal STM, and verb retrieval when compared to preoperative scores.

Second, consistent with our predictions, tumour volume was a significant determinant of the degree of improvement/decline. However, the directionality of this effect was inconsistent with our predictions – that is, larger volumes were actually found to be significantly predictive of greater decline across the pre-operative and follow-up period. Further, inconsistent with our predictions, tumour grade and type were not significant determinants of the degree of improvement/decline. However, a significant effect was obtained for age, with greater age associated with greater improvement between the preoperative and follow-up period on the core skill measure of Auditory Word Recognition. This finding was inconsistent with our original predictions, in which we predicted that older age would be predictive of greater decline, and less improvement across the surgical phases. It may be the case that this significant finding is due to younger patients showing ceiling effects during the preoperative phase, which limited their potential to show follow-up improvement.

Third, consistent with our predictions regarding tumour localisation, during preoperative testing, those with a left temporal tumour were significantly more impaired in *accessing semantic knowledge* than the other groups. However, none of our other predictions regarding tumour localisation were supported.

This study tested patients at three timepoints: preoperatively, immediately postoperatively, and at least three months following surgery. However, scores at immediate postoperative testing were highly variable across patients and did not appear to be useful predictors of post-acute performance (see Bello et al., 2007; Ilmberger et al., 2008; Santini et al., 2012; Teixidor, Gatignol, Leroy, Masuet-Aumatell, Capelle, & Duffau, 2007, for similar findings). This variability in outcomes is perhaps not surprising when we consider the various factors that are likely to influence performance at this time point. Some of these factors may lead to improvements in performance, such as the reduction in tumour mass following resection, which may enhance performance, whilst other factors may operate to impede performance at this timepoint – for example, the effects of medication, anaesthesia, and post-operative inflammation and swelling (Heimans & Reijneveld, 2012). For these reasons, we will not consider immediate postoperative performance in any further detail here.

The following section will explore the sensitivity of each of our measures, followed by a discussion regarding the key findings and predictions.

Implications for assessment of language. Whilst the focus on language outcomes was primarily based on long-term follow-up, it is nonetheless useful to examine the overall incidence of pre-operative impairment – and how this compares to previous studies – to gain better insights into the sensitivity of the BLAST protocol as a whole. Our results confirm that, when compared to more conventional language assessments, the BLAST appears to be particularly sensitive to deficits. Pre-operatively, a surprisingly high 95% of patients exhibited significant impairment on at least one BLAST subtest, whilst 71% were impaired on at least one skill score. These rates are surprisingly high when compared to other figures previously reported, which range between 37% and 64% (Bello et al., 2007; Haglund et al., 1994; Recht et al., 1989; Sanai et al., 2008; Tandon & Mahapatra, 1993; Whittle et al., 1998). The rates we obtained here are perhaps most comparable to those of studies that have utilised a neuropsychological approach. For example, Papagno and colleagues (2012) assessed language functions in a sample of high and low-grade tumour patients using a range of tasks, including verbal fluency, real-word and non-word sentence repetition, and naming of famous faces, pictured objects, and actions. Overall, rates of pre-operative impairment ranged between 60 and 79 percent. Further, in a prospective study involving 29 glioma patients (HGG:17, LLG: 12), Talacchi and colleagues (2011) administered a number of tasks, including word fluency (FAS), trail making, verbal digit span, immediate and delayed word recall, visual object naming, and copy design. During the pre-operative phase, 79% of

patients showed impairment on at least one task (Talacchi et al., 2011).

Of course, the absolute incidence of impairment observed in any cohort will depend as much upon the individuals tested as it does the methods of testing. Our sample included individuals with a wide range of tumour types and locations, whereas several previous studies have focused on low-grade tumours specifically (e.g., Duffau et al., 2008; Satoer et al., 2013; 2014; Teixidor et al., 2007). Importantly however, language and cognitive deficits have been found to be more marked in patients with high-grade glioma, and malignant tumours (Duffau, 2005; Heimans & Reijneveld, 2012; Noll, 2014; Hom & Reitan, 1984; Imperato et al., 1990; Klein & Heimans, 2004; Miotto et al., 2011; Tucha, Smely, Preier & Lange, 2000; Whittle et al., 1998). Consequently, our inclusion of these latter cases may have increased the overall incidence of language impairments in our sample. Conversely, many previous studies have restricted inclusion to those with left hemisphere tumours (e.g., Ilmberger et al., 2008; Santini et al., 2012; Satoer et al., 2013). Restricting the sample in this way may operate to increase the incidence of significant impairment; importantly however, these studies have predominately used more conventional assessments to measure language functioning, which may not be sensitive to the often subtle language deficits that occur in neurological tumour populations.

The primary aim of the BLAST is not to assess overall performance on any one task, but rather to generate numerical estimates of various hypothetical language processes. Since no previous studies have included these sorts of measures, we cannot compare our results from the skills measures directly with any previous studies.

Turning now to the specific tasks we used, the four subtests that were most consistent at detecting accuracy impairments across the testing phases were: *Picture-Word Verification, Verb Generation, Picture Naming,* and *Letter Fluency. Non-Word Repetition* was the least sensitive. These findings are consistent with a number of studies that report picture naming and fluency measures are amongst the most sensitive at detecting significant impairment in neurological tumour populations (e.g. Faulkner et al., 2017; Papagno et al., 2012; Santini et al., 2012). Possibly, this finding may be attributed to an extensive wide-spread network of cortical structures, involving left anterior, temporal and posterior regions that are engaged when naming a picture (see esp. Baldo et al., 2013; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Hillis et al., 2006). From a clinical perspective, the identification of tasks that are sensitive during both the acute surgical phase and long-term follow-up is important, as it highlights the most

appropriate tasks to administer in this population. This not only fulfils an important goal of task sensitivity and brevity, but also ensures that patients are not administered tasks unnecessarily.

With respect to the core skills, the most sensitive measures were *verb retrieval, accessing semantic knowledge, lexical selection,* and *verbal short-term memory*. Conversely, *Phonological encoding* was the least sensitive core skill, with no patients exhibiting significant impairment across both the pre-operative and follow-up phase. This is not surprising given the subtest that contributed to this skill - non-word repetition - was the least sensitive at the subtest level. Further, as shown in Figures 2.4 and 2.5, there was very little tumour overlap between patients in the regions hypothesised to be implicated in phonological encoding (left temporal and left parietal). Consequently, it may be the case that our sample was non-representative, wherein we simply did not have any patients who had lesions in the region that we are hypothesising.

Long-term outcomes. We now consider the issue of performance change between the preoperative and the follow-up phases. In line with our predictions, we found that a high proportion of patients showed long-term deficits that persisted beyond the acute surgical phase. Specifically, 78 percent of patients were significantly impaired on at least one subtest during the follow-up period, compared to 95 percent and 92 percent at the pre-operative and post-operative phases respectively. A similar pattern was evident at the core skills level, with 49 percent of patients impaired on at least one core skill during the follow-up period compared to 71 percent pre-operatively. The rates of long-term impairment in the current study are on the high side, when compared to previous studies. For example, in studies that use more conventional measures of aphasia such as aphasia quotients, the incidence of impairment at long-term follow-up generally does not exceed 20% (e.g. Duffau et al., 2008; Ilmberger et al., 2008; Sanai et al., 2008). However, even those that define impairment as below-normal performance on one or more neuropsychological language tasks typically report lower incidences of impairment than we obtained in the current study; impairment rates for these previous studies generally do not exceed 50 percent (e.g. Papagno et al., 2012). Consequently, the BLAST can be considered a highly sensitive method of detecting significant language difficulty in a neurological tumour population. Of course, given that our preoperative rates of impairment were generally higher when compared to previous studies, we would also expect a higher incidence of significant impairment at long-term follow-up.

Nonetheless, our findings lend support to the use of neuropsychological approaches to assess language functions.

One reason for the high incidence of impairment in our assessment is that many of the tasks make high demands on control processes. For example, tasks such as Verb Generation, Stroop, and Letter and Category Fluency are likely to require a number of higher-level cognitive processes that may extend well beyond the verbal domain (e.g., processing speed, performance monitoring, vigilance, sustained attention and maintenance, inhibition, and internally-driven generation; Alexander, 2006; Stuss et al., 1998). Consider the verbal fluency task. Successful completion of this task requires individuals to utilise task-setting by formulating a response goal, whilst also simultaneously energising their schema and maintaining an activated state throughout the task (Alexander, 2006). Individuals must also suppress any irrelevant or competing schema and intrusions, and ensure that responses are constantly monitored so that inappropriate or perseverative errors are not produced (Alexander, 2006; Shao et al., 2014). Such higher-level skills have typically been localised to right frontal, dorsolateral and superior medial regions of the prefrontal cortex (Stuss et al., 1995; Ridderinkhof et al., 2004). Consequently, low scores on some of our tasks might not reflect poor language abilities per se. We return to this issue in Chapter 5.

Importantly though, one advantage of the core skill measures is that many are based on comparisons between two or more measures. Therefore, these scores may factor out some of the more general cognitive factors that are likely to contribute substantially to raw task scores. Indeed, this was evident when we compare the rates of significant impairment between the subtest and core skills level. For example, 49% of patients showed significant impairment on at least one skill measure during long-term follow-up compared to a substantially higher 78% at the BLAST subtest level. Importantly, one of the specific goals of this current thesis is to explore in detail the relationship between language-specific and more domain-general cognitive skills. The results of this exploration will be presented in Chapter 5.

Another reason for the discrepancy between the current high rates of impairments and those reported previously may be due to differences in the pathological characteristics of the patient samples. Many previous studies that have examined long-term outcomes in tumour populations have predominantly assessed only glioma patients (most often low-grade glioma; e.g. Duffau et al., 2008; Satoer et al., 2013; 2014; Teixidor et al., 2007). Specifically, in the aforementioned sample of 115 individuals with low-grade glioma, Duffau and colleagues

(2008) reported a complete recovery to baseline performance in 98 percent of patients three months' post-surgery. Similar observations have been reported by Teixidor and colleagues (2007), who concluded that tumour patients not only recovered language and cognitive functions, but may also improve upon pre-operative status. However, similar to Papagno and colleagues (2012), the current sample included a diverse range of tumour types, with a number being high-grade and malignant glioblastoma. This is an important distinction that may contribute to the differential rates of impairment, given high-grade tumours are often associated with greater rates of neurocognitive sequela relative to lower-grade lesions. Indeed, in a recent prospective study involving 32 patients with high-grade glioma, patients experienced deterioration in attention, information processing and psychomotor speed during the eight and sixteen-month follow-up assessments compared to baseline performance between surgery and radiation (Bosma et al., 2007; see also Corn et al., 2009).

Whilst the high incidence of long-term impairment was consistent with our original predictions, there was also evidence of a general improvement across the acute and follow-up phases. Specifically, 57 percent of patients showed a substantial improvement in scores on at least one BLAST subtask, and 22 percent improved substantially on at least one core skill. Such patterns are consistent with a number of previous studies that report a long-term improvement in language performance within tumour populations (e.g. Duffau et al., 2008; Sanai et al., 2008; Santini et al., 2012; Teixidor et al., 2007); this may be attributed to the beneficial effects of tumour resection, and possibly also the functional reorganisation that occurs following surgery. Indeed, a number of studies have found that surgical resection can reduce pressure on the surrounding neural regions proximate to the tumour, resulting in some immediate improvement to functionality. For example, in an aforementioned study involving 29 glioma patients, Talacchi and colleagues (2011) reported that 24% of patients exhibited an immediate post-surgical improvement, when compared to pre-operative performance. However, it remains unclear whether the recovery observed in the current study is due to a release of pressure on the surrounding regions that were previously implicated in that skill, or instead whether the loss of tumour bulk operates to facilitate functional reorganisation.

For many key measures, preoperative scores were a significant predictor of performance at long-term follow-up. At the individual task level, this was the case for *Category Fluency*. However, at the core skills level, it was the case for *accessing semantic knowledge*, *lexical selection*, *phonological encoding*, and *verbal short-term memory*. These findings are perhaps not surprising, given that those who were impaired preoperatively were

also those most likely to have residual difficulties at long-term follow-up. It is interesting to note that the core skills measures yielded the most consistent outcomes here; this suggests that a core skills approach may provide more reliable and robust measures of language function than raw performance skills on specific tasks.

We also investigated whether the following factors influenced long-term language outcomes: *patient age, tumour volume (cm^3) and tumour type*. Despite our predictions that greater age would be a significant predictor of less improvement, and even greater decline, at follow-up, the opposite was found. Specifically, on our core skill measure of auditory word recognition, greater age was actually predictive of greater *improvement* between the pre-operative and follow-up testing phases. This result likely reflects younger patients showing a ceiling effect during the preoperative phase, which limited their potential to show substantial improvement at follow-up. Despite our original predictions, no other significant age effects were observed. Of this, the absence of findings may be due to our small sample size and a lack of power to detect significant effects. Such findings may also be due to other factors confounded with age. For example, patients in our 30 - 50 year age group tended to have larger lesion volumes, on average that the older patients; as discussed below, larger lesion volumes were associated with greater deterioration across the acute pre-operative period and follow-up. To overcome the confound between age and lesion size, future studies may wish to factor out lesion size using multiple regression. For the purpose of the current study, the sample size would simply be too small to detect any reliable and valid effects.

Our second hypothesis was that greater tumour volume (cm^3) would be associated with greater improvement across the pre-operative and follow-up period. Inconsistent with this prediction, we found that larger tumour volumes (cm^3) were actually associated with significantly *less* improvement across the pre-operative and follow-up period on the following core skills: *accessing semantic knowledge* and *articulatory motor planning*. Specifically, for the core skill of *articulatory motor planning*, our three patients with the largest lesions volumes[13] (C.A., A.V.G., & P.C.) exhibited a general decline in scores from the pre-operative to follow-up period, whilst those with the smallest lesion volume (K.W., L.W., & N.O.H.) all showed a general pattern of improvement. Importantly, those with large lesion volumes all presented with low-grade lesions, and had not received prior

---

[13] Our patients with the largest and smallest lesion volumes, TD and RS respectively, were excluded from this analysis as they did not complete the pre-operative articulation subtest.

chemotherapy or radiation, nor were currently taking antiepileptic drugs, so this result is unlikely to be driven by the confounding effects of some other variable. Moreover, given low-grade tumours are more insidious and slow growing in nature (see Duffau, 2005), it is likely that larger tumours have been present for a very long time in these individuals. Evidence suggests that in patients with slow-growing progressive tumours, there is likely to be a great deal of functional reorganisation as the tumour grows (Duffau, 2005). Accordingly, it is possible that removal of the tumour may provide only limited opportunities for regain of function. Another possibility is that a large, slow growing tumour may impact more severely on surrounding neuronal tissue than a smaller one, to the point where it cannot completely recover its function substantially, even after the tumour is removed. Moreover, it is possible that pre-surgical oedema contributed to the significant effect between lesion volume and long-term outcomes, which was not accounted for in the current study.

Another possibility that may account for the above finding is the extent of resection (EOR). For example, it is entirely possible that larger tumours undergo a greater extent of surgical resection compared to smaller tumours; this may have some functional impact on the surrounding neural regions. Investigating the effects of tumour volume on long-term outcomes appears to be an area where research is lacking, and thus requires a systematic exploration in future studies. Nonetheless, from a clinical perspective, this finding may suggest that lesion volume can serve to mediate long-term outcomes, whilst also highlighting that patients with larger lesion volumes on pre-operative scans may be ideal candidates for close post-surgical monitoring and rehabilitation. However, any interpretations should be tentative given our small sample size and the variable time in which the follow-up assessments were conducted, ranging from three to 12.2 months (average time since surgery = 5.7 months). We also cannot rule out the possibility that in our sample, lesion size and location were confounded in some way; it would be difficult to systematically study this hypothesis in such a small and highly variable sample.

Our next hypothesis was that higher-grade tumours would result in disproportionately less improvement and/or greater decline in core-skills scores across testing phases. These predictions were based on previous findings that have found greater incidence of neurocognitive decline in patients with higher-grade tumours (e.g. Whittle et al., 1998). However, this hypothesis was not supported. This may possibly be due to a floor effect; those with high-grade lesions scored poorly on many measures pre-operatively, so may have had limited opportunities to show further decline at long-term follow-up. Selection bias may also

have mitigated any effects here; those participants who consented to being tested at long-term follow-up were generally those whose medical status had not deteriorated substantially since surgery. For example, a large portion (37%) of eligible patients with high-grade tumours passed away before the follow-up assessment could occur. Therefore, the "true" incidence of long-term decline may be higher. Importantly however, exploratory analyses revealed very few significant differences between those who passed away during the testing phases and those included in the current sample. The only exception to this was *goal-driven response selection*, with those who passed away scoring significantly more poorly than those who did not; this effect was not driven by tumour volume, tumour location or participant age. One explanation for this finding may be our broad definition of goal-driven response selection. That is, we assumed the existence of a very generalised "control" function that is involved not only in the suppression of well-learned responses but also the resolution of conflict. These types of functions may utilise more generalised cognitive resources compared to our remaining core skills measures and are therefore influenced by more generic cognitive functions (see Alexander, 2006; Novick et al, 2010). In other words, the above finding may not be specific to the core skill of goal-driven response selection, but may instead reflect the involvement of broader cognitive processes. This issue is explored in greater detail in Faulkner & Wilshire (in press, 2018).

Another possible confounding factor here is seizure status: a high proportion of low-grade tumour patients (30%) were receiving anti-epileptic medications at the time of follow-up, indicating a higher epilepsy burden than for high-grade patients. This is an important observation as epileptic seizures and anti-epileptic drugs (AED's) have long been associated with a range of neurocognitive and motor deficits, including visual disturbances, hearing loss, extremity weakness, and cognitive slowing (Bosma et al., 2007; Klein et al., 2003; Lieu, & Howng, 2000; Mckee, Blacklaw, Butler, Gillham, & Brodie, 1992; Whittle, Smith, Navoo, & Collie, 2004). To overcome the potential confound between performance and epilepsy burden, future studies may wish to factor out medication using multiple regression. Again, however, for the purpose of the current study, the sample size would simply be too small to detect any reliable and valid effects.

Anatomical Specialisation. Our final objective was to examine the specific anatomical predictions associated with the BLAST subtests and core skills – irrespective of time of testing. Based on previous literature (see Faulkner et al., 2017), we made the following predictions regarding each of our core skills: patients with left frontal lesions will

have significantly lower average scores on *articulatory motor planning, goal-driven response selection* and *verb retrieval.* Patients with left temporal lesions will have significantly lower average scores on *accessing semantic knowledge, lexical selection, auditory word recognition, verbal short-term memory,* and *phonological encoding.* Finally, patients with left parietal lesions will have significantly lower average scores on *verbal short-term memory* and *phonological encoding.*

However, there was no anatomical specificity observed. A number of factors may have contributed to our findings. These include the small sample size, which limits the statistical power to detect significant effects, and the effects of other potentially confounding variables, such as the length of time since surgery, tumour volume and age. In order to further investigate the effects of tumour localisation on our skill measures, future studies should endeavour to conduct multiple regression to counter some of these confounding variables. Again, for the purpose of the current study, the sample size was simply too small to detect any reliable and valid effects. Further, as shown in Figure 2.2, there was very little overlap of tumours in our sample. Accordingly, our sample may not have representative cases to reliably detect significant localisation effects of the skills we were hypothesising.

Conclusion. The findings presented in this chapter suggest that pre-operative assessments using a sensitive core skills approach has the capacity to provide a rich source of information regarding the likely long-term implications of surgery. This information can be used to predict long-term outcomes in addition to guiding pre and post-surgical planning and long-term interventions. Our approach may offer benefits over protocols that are largely developed for and validated on stroke populations. Further, using a sensitive assessment tool developed specifically for tumour patients, the current investigation has demonstrated that the likelihood of language recovery following tumour surgery may be lower than has been suggested in previous studies that use more conventional language measures. Indeed, many patients who are significantly impaired during the acute surgical phase are likely to remain impaired during long-term follow-up. Moreover, in some cases, we see a significant decline in performance during follow-up testing. One factor that may influence long-term outcomes is lesion volumes, with larger lesions (cm^3) associated with *less* long-term improvement across the acute surgical and follow-up periods. Accordingly, this observation may highlight those patients who are ideal candidates for post-operative rehabilitation.

Chapter 3: Core Skills and Complex Language

Introduction

The "core" cognitive skills identified in the BLAST are all based on tasks involving the production or comprehension of single words. Clearly, effective communication requires much more than single words – it requires successful performance at the sentence level and also at the discourse level (Satoer et al., 2013). An avenue that has so far received little attention is the relationship between our core skills measures and more complex aspects of sentence-level comprehension and production. Assessing this relationship can provide valuable insights into the nature in which language can be decomposed into more elemental cognitive skills. A thorough exploration of sentence-level language is beyond the scope of this thesis, however, the following sections will explore the theory and evidence relating to sentence-level speech production and comprehension, with particular attention given to accounts that relate to the core skills within the BLAST. Finally, the current study and hypotheses will be discussed.

Sentence-Level Language Production

The ability to formulate coherent multi-word utterances is likely to require a number of processes. There are wide theoretical disagreements as to the processes that are engaged during sentence-level production (see especially Dell, 1986; Garrett, 1975; Levelt, 1989; Levelt, 1999; Stemberger, 1985). However, most models agree that, at a very general level, producing connected speech involves three broad types of processing: the *conceptualisation* of abstract representations, the *formulation* of sentence plans, and the *execution* of language components (Alario & Cohen, 2004; Bock & Levelt, 1994; Levelt, 1989). *Conceptualisation* is considered to be a pre-linguistic process. It involves developing a conceptual message intention – in the form of a concept or proposition - which is then used to drive the formulation of the sentence. *Formulation* involves specifying the key lexical elements and the overall structure of the sentence, and drawing on appropriate grammatical concepts. *Execution* refers to the articulatory planning and execution involved in actually producing the sentence (Alario & Cohen, 2004; Harley, 2001; Levelt, 1989). Our discussion will be focussed on those processes that come under the general rubric of formulation.

Some theories of sentence-level production propose that sentence formulation is achieved through a frame allocation process. For example, according to Garrett's model of sentence-production, sentence planning begins with the development of an abstract, non-

linguistic representation of the sentence, which specifies the concepts to be described and their interrelationships (Garrett, 1975; 1976; 1980). Then, two different types of operations take place in succession. First at the *functional* stage, the lexical labels for all substantive words (open-class words: nouns, verbs, adjectives, etc.) are retrieved, and the overall grammatical relationships amongst these words are established (see also Alario & Cohen, 2004; Biassou, Obler, Nespoulous, Dordain, & Harris, 1997). These types of word items are then inserted into an appropriate syntactic frame, resulting in a fully formed plan of the sentence (Garrett, 1975; 1976; 1980; see also Harley, 2001). Following this, the appropriate positions of the syntactic frame need to be specified. This occurs at the *positional* stage, wherein the appropriate closed-class words are retrieved (e.g., propositions, determiners and conjunctions) as required by the grammatical structure of the sentence and by more general grammatical rules.

This early model was supported by evidence of a double dissociation between open class word production (substantive words, such as nouns, verbs), and closed class word production (determiner, conjunctions etc.). In non-fluent aphasia, open class words are often well represented in connected speech, but obligatory closed class words (e.g., articles, prepositions) are often absent (Goodglass et al., 2001; Rochon, Saffran, Berndt, & Schwartz, 2000; Saffran, Berndt, & Schwartz, 1989). This pattern of speech is often referred to as *agrammatic*, because the speech appears to lack a normal grammatical structure. For example, when asked to describe a picture of a woman doing the dishes with her children around her, an agrammatic patient may present with a simplified formation of the substantive words in the sentence, such as "woman, dishes, children, water", however, the grammatical structure of the narrative is usually omitted. Indeed, this agrammatic pattern of speech that is often seen in non-fluent aphasia has been the subject of much study in its own right, because of its potential to shed light on the process of grammatical role assignment in sentence formulation (see for example, Bastiaanse & Jonkers, 2012, for a recent review). Conversely, in Wernicke's aphasia, closed class vocabulary is usually well represented in speech, but open class vocabulary is frequently missing or incorrect (Alario & Cohen, 2004; Berndt, 2001; Goodglass et al., 2001). Such patterns provide support for two different types of processes involved in the formulation and planning of sentences.

Some recent theories of sentence planning place particular emphasis on the selection and retrieval of the verb, which is not only an essential component for a sentence, but also constrains its grammatical structure (Marshall et al., 1998; see also Ahrens, 2003; Berndt et

al., 1997; Levelt 1989; 1999; Mätzig Druks, Masterson, & Vigliocco, 2009; Pickering & Branigan, 1998; Shapiro & Levine, 1990; Trueswell & Kim, 1998). In fact, verbs play a key role at both the conceptual and the grammatical formulation levels. At the conceptual level, verbs are distinct from other word classes as they act to specify the relationships between nouns and their thematic roles in a sentence (Marshall et al., 1998; Rofes and Miceli 2014); at the syntactic level, verbs are necessary for specifying the direct and indirect objects of a sentence in order to formulate an argument structure (Grimshaw, 2000). Consider the example of "*ran*", which requires no direct or indirect object to formulate a connected and meaningful sentence (e.g. "John ran"), whereas "gave" requires both a direct and an indirect object (e.g. "John *gave* the present to Mary"). Consequently, difficulty retrieving verbs is likely to affect the construction of sentence-level speech.

There is some evidence to support this claim. For example, Berndt, Burton, Haendiges, and Mitchum (2002) reported that patients who demonstrated verb deficits on a picture naming task also showed impaired production of sentences. Moreover, individuals with non-fluent aphasia, who produce an agrammatic pattern of speech, often perform particularly poorly on tasks that place demands on verb retrieval. For example, in an earlier study, Zingeser and Berndt (1990) administered a picture naming and action description task to five patients with agrammatism and five patients with anomic aphasia. Those with agrammatism produced significantly fewer verbs than nouns, relative to the anomic patients (see also Mätzig et al., 2009; Marshall et al., 1998). It has been suggested that this difficulty with verbs may underpin the severe sentence production problems often observed in agrammatic non-fluent aphasia (e.g. Marshall et al., 1998). Indeed, there is some evidence that in non-fluent aphasia, poor scores on verb production tasks, such as action naming tasks, are associated with poor accuracy in sentence production, particularly on sentences that contain multiple argument structures (Marshall et al., 1998; Thompson, Lange, Schneider, & Shapiro, 1997).

In the last decade or so, there has been particular interest in the relationship between the conceptualisation and formulation stages. In spontaneous speech, we not only need to create an internally generated conceptual message, but also, this message needs to be able to take control of later stages of the sentence planning process. This control needs to be powerful enough to ensure that other words activated in the mind of the speaker do not take control of the system and become erroneously produced. This ability is sometimes measured using unconstrained tasks, such as those requiring the generation of complex sentences, or the

spontaneous production of connected speech. A clinical condition suggestive of a failure at this conceptual control level is Transcortical Motor Aphasia. Patients with this condition rarely initiate spontaneous speech; however, when they are provided with a specific task or message to convey (such as a pictured scene to describe) they may produce entirely well-formed speech (Alexander & Schmitt, 1980; Chapados & Petrides, 2013; de Lacy Costello and Warrington, 1989; Freedman et al., 1984; Goldberg, Mayer, & Toglia, 1981; Pai, 1999; Ziegler et al., 1997). Such individuals also tend to perform poorly on word generation tasks such as the verbal fluency task, which would also be consistent with an inability to generate concepts spontaneously, with minimal guidance as to their nature.

A related skill in sentence-level production is the ability to manage the level of activity in the lexicon (Scott & Wilshire, 2010). As discussed in Chapter 1, it has been proposed that language representations become activated via a process of spreading activation (see Chapter 1). Each linguistic unit that is activated will in turn activate those that share some of these components (e.g., the lexical unit for cat, once activated, will spread activation to the unit for dog via its shared semantic interconnections; see Dell, 1986; Levelt, 1999). Therefore, the ability to speak efficiently also involves managing activation levels within the lexicon so that words that are activated, but not planned for the upcoming sentence, do not become erroneously selected in place of the correct words. The challenge is all the greater when producing a longer sentence, because the speaker may have in mind a number of concepts and other elements related to the sentence being planned, but must ensure that each of these only reaches full activation when it is required for production. Several theorists have proposed that a dedicated control system is called upon to manage this type of noisy competition (Robinson, Shallice, & Cipolotti, 2005), which functions to oversee the flow of activation and resolve conflict within the lexicon. A number of suggestions have been put forth to account for the mechanisms that underpin this process, with one suggesting competition is managed by inhibiting previously activated lexical items (Biegler, Crowther, & Martin, 2008). Another position proposes this control system operates through a biasing process that can either actively enhance or inhibit representations in the lexicon, depending upon the current production goal (Hamilton & Martin, 2005; January, Trueswell, & Thompson-Schill, 2009; Schnur, Schwartz, Brecher, & Hodgson, 2006; Thompson-Schill et al., 1997; Wilshire & McCarthy, 2002).

Further, some individuals with non-fluent aphasia exhibit delays producing words under conditions where the words are likely to be highly active and in competition with the

desired item (Novick, Kan & Trueswell, 2009; Schnur et al., 2006; Scott & Wilshire, 2010). For example, these individuals may have particular difficulty inhibiting the inappropriate colour responses on the Stroop task (Hamilton & Martin, 2005; Scott & Wilshire, 2010), and may have difficulty with tasks that involve repeated naming of a small set of closely related semantic items. Both these kinds of tasks would seem to create a high level of competition amongst the concept/lexical item to be produced and other task-irrelevant concepts/lexical items (Novick et al., 2009; Robinson et al., 2005; Thompson-Schill et al., 1997; Thompson-Schill et al., 1999). Another task that has been argued to have a strong competition resolution component is the verb generation task (see Novick et al., 2009; Thompson-Schill et al., 1999). Indeed, recent studies have reported that individuals with non-fluent aphasia perform disproportionately poorly on this task when the stimulus noun is associated with multiple correct alternatives (e.g. baby -> "*cry*", "*crawl*", "*sleep*") compared to those where there is one dominant response (e.g. scissors -> "*cut*") (e.g., Cameron Jones, 2008; Novick et al., 2009; Robinson et al., 2005). Some researchers have suggested that in the former condition, there is greater competition between multiple alternatives (Novick et al., 2009; Thompson-Schill et al., 1999). Accordingly, impairments resolving competition at the lexical stage is likely to impact on speech at the sentence-level.

Sentence-Level Speech Comprehension

Turning now to theories of sentence-level speech comprehension, a large body of evidence suggests that the listener makes use of multiple sources of information to understand sentence meaning; this includes not only the grammatical cues contained within the sentence itself, but also broader contextual information about the world and the nature of the current discourse (Bates & MacWhinney, 1989; Clifton, 1993; Jurafsky, 1996; Daneman & Merikle, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994; Spivey-Knowlton, Trueswell & Tanenhaus, 1993). However, for more complex sentences, broader contextual meaning alone is usually insufficient to establish the correct semantic relationships amongst the sentence elements. In this situation, an individual must retain and manage those sentence elements, so that they can be appropriately conceptually integrated across the entire sentence (Altmann & Kamide, 1999; Novick et al., 2009). At the same time, they need to decode the incoming speech stream with close to real-time speed, so that no sentence elements are missed. Indeed, a number of sources of evidence suggest that elements of sentences are assigned meaning as they are encountered, and often before the speaker has concluded the sentence (e.g., Altmann & Kamide, 1999; Ferreira & Clifton, 1986; Novick et al., 2009;

Trueswell, Tanenhaus, & Garnsey, 1994; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). This type of incremental processing means that later information in the sentence requires the listener to revise and update their interpretation of previous elements (e.g., *I saw a man eating shark at a restaurant*). Accordingly, this is a demanding and complex task.

Many current theories of sentence-level language comprehension emphasise the interactive nature of the comprehension process, and also the inherent ambiguity of linguistic processing. One prominent class of theories proposes that listeners utilise multiple sources of syntactic, semantic and discourse information, called *constraints*, in order to rapidly disambiguate an acoustic speech signal (see Boland, Tanenhaus, & Garnsey, 1990; MacDonald, 1994; MacDonald, Pearlmutter, & Seidenberg, 1994). In these models, any ambiguity between these sources is resolved through competition (MacDonald, 1994; MacDonald et al., 1994). The constraint that receives the greatest activation determines the construction of a sentence. Consequently, a failure to resolve such competition may result in severe processing difficulties during sentence-level speech comprehension.

A number of studies have suggested that left frontal regions, including the Broca's area, may be involved in the syntactic processing of complex, temporally ambiguous sentences (e.g. see Novick et al., 2009; Stromswold, Caplan, Alpert, & Rauch, 1996). For example, in a recent study, Novick, Kan, Trueswell and Thompson-Schill (2009) investigated language processes, including sentence comprehension, in one patient, I.G., who presented with a circumscribed lesion to the left inferior frontal gyrus. The sentence comprehension task involved responding to two types of sentences - one that was unambiguous (e.g. "*put the frog that's on the napkin into the box*") and one that was temporally ambiguous (e.g. "*put the frog on the napkin into the box*"). Overall, I.G. exhibited significant difficulty over-riding and resolving his initial interpretation when temporally ambiguous sentences were produced. This difficulty resolving conflict was also evident at the individual word level as displayed in a verbal fluency paradigm (see Novick et al., 2009).

Prominent theories of sentence comprehension also emphasise the important role of verbal short-term memory. Indeed, several studies report an association between comprehension capacity and the ability to retain a sequence of phonological/lexical representations (e.g. Carretti, Borella, Cornoldi, & De Beni, 2009; Gernsbacher, Varner, & Faust, 1990). Verbal short-term memory is classically conceptualised as a limited capacity storage system that stores and maintains recently encountered phonological strings (see e.g.,

Baddeley, 1984). It is proposed that the phonological information held in this store undergoes rapid decay unless actively rehearsed. The effectiveness of verbal short-term memory has traditionally been evaluated by tasks that involve the immediate verbatim recall of some sort of phonological sequence, such as a series of digits or a non-word (Jacquemot & Scott, 2006; Romero Lauro, Reis, Cohen, Cecchetto & Papagno, 2010). As discussed in Chapter 1, the repetition of non-words is held to place heavy demands on a short-term storage site because unfamiliar phonological items are not stored in our mental lexicon (Archibald & Gathercole, 2007; Martin & Gupta, 2004). Individuals with brain damage or dysfunction may perform poorly on these tasks, suggesting a specific impairment involving the phonological short-term store (Martin & Saffron, 1997; Vallarb, Di Bettac & Silveri, 1997). Indeed, in a recent voxel-based lesion symptom mapping (VLSM) study, Baldo, Katseff and Dronkers (2012) reported neural overlap between digit span and repetition tasks; specifically, the left posterior temporo-parietal region was critical for both tasks, suggesting span and repetition abilities are underpinned by the same mechanisms.

It has been suggested that this phonologically-based verbal short-term store is called upon when the sentence being processed is particularly long and/or complex. Accordingly, deficits involving this short-term store will result in difficulties comprehending some sentences (Daneman & Merikle, 1996; Saffran, 1990). There are some studies that have looked at the neural substrates of sentence comprehension deficits, and they find the same areas implicated as the ones in the phonological store impairments, suggesting the two capacities are indeed linked (Gvion, & Friedmann, 2012; Papagno, Cecchetto, Reati, & Bello, 2007). The current study provides an opportunity to examine this relationship in a neurological tumour population, which to our knowledge has been limited in the literature.

Of course, sentence comprehension also involves additional processes not discussed here. Some researchers have suggested it also relies on a special form of working memory called semantic short-term memory, in which lexical elements in the phrase currently being processed are held in their particular order until they can be integrated into a phrasal conceptual proposition (Biegler et al., 2008). This form of working memory is considered separate and distinct from verbal (phonological) short-term memory. Also, effective comprehension likely relies on a range of higher level skills that support the integration of newly encountered information with previously processed material (Daneman & Merikle, 1996; Caplan & Waters, 1999; Caspari, Parkinson, LaPointe, & Katz, 1998). These types of processes are beyond the scope of the present review.

The Current Study: Task Selection and Hypotheses

The goal of this Chapter is to investigate the relationship between sentence-level comprehension and production and the core skills approach utilised in the BLAST. Addressing this relationship can provide insights into the nature in which sentence-level speech can be decomposed into simpler, core skills. To address this question, sentence-level speech production will be measured using the Quantitative Production Analysis (QPA) (Saffran, Berndt, & Schwartz, 1989). In this assessment, the participant is asked to retell a familiar short story (e.g., Cinderella), and then various different elements of speech output are measured or counted. This test was selected for the current study as it has been well validated for use with stroke populations, and also, it generates specific quantitative measures for different aspects of speech production (e.g., rate of speech, the proportion of closed-class words, and the extent to which an individual 'struggles' to produce a coherent narrative'). Some specific measures from the QPA that we will use here include: *closed class ratio* (a measure of the proportion of closed class words in the narrative); *speech rate* (number of words produced per minute); and the s*truggle measure,* which indexes the proportion of words produced that contribute to the speaker's narrative (a low score on this measure suggests the speaker is having difficulty generating a coherent utterance). These latter two measures were selected as they may prove to be good indicators of deficits in the kind of goal-driven language generation abilities discussed above.

To assess sentence-level comprehension, we selected the Test for the Reception of Grammar (TROG) (Bishop, 1989). This task was selected as it serves as a well-validated measure of sentence-level speech comprehension, and provides separate measures for vocabulary comprehension and also for the comprehension of grammatical relationships within sentences. Further, the TROG examines a range of syntactic constructs, which range from single-words to simple subject-verb-object (SVO) sentences (e.g., "*the cat chases the dog*"), to more complex sentence constructions (e.g. "*the cat the cow chases is black*" "*the boy the dog chases is big*") (Martin & Romani, 1994; Martin & Feher, 1990; Vallar & Baddeley, 1984). The task also includes examples of constructions in which early ambiguity is not resolved until later in the sentence (e.g. "*The pencil is on the book that is yellow*", "*the square is in the star that is blue*"); this is a sentence type that has been found to be particularly challenging for individuals with a deficit in phonological short-term memory (Gvion & Friedman, 2012), and may also be a good indicator of deficits in goal-driven

language capacities as the individual is required to overcome their initial processing until the sentence is complete (MacDonald et al., 1994).

The TROG generates a number of different indices related to sentence comprehension abilities. It generates a "*Total blocks passed*" measure, which provides an index of overall sentence comprehension accuracy. A "block" refers to a set of sentences with similar priorities. Blocks are presented in order of difficulty, commencing with very simple structures (e.g., nouns, verbs), and terminating with complex structures (e.g., embedded sentence structures). The TROG also generates specific measures relating to the total number of lexical and grammatical errors on various blocks.

Testing time is limited during the acute preoperative and postoperative phases, so it was not possible to examine these additional measures at these timepoints. Therefore, all the measures reported here were obtained during the long-term follow-up testing phase (average time since surgery = 5.7 months), including the BLAST measures to which they were compared.

Our predictions with respect to the QPA were as follows. First, given that verb retrieval is thought to be important for guiding the construction of a sentence, we predicted that poor scores on the BLAST verb retrieval measure would be associated with difficulty forming grammatically correct sentences. Consequently, we would predict poor verb retrieval scores to be associated with lower scores on the QPA closed class ratio measure (which is likely to be an index of grammatical well-formedness). These difficulties forming a coherent utterance are also likely to result in a lower rate of speech (i.e., lower scores on the QPA speech rate measure), and subsequently, lower overall scores on the QPA struggle measure.

Second, the BLAST measure of goal-driven response selection indexes an individual's capacity to generate speech based on an internally represented goal. We have argued above that one important aspect of this skill is the ability to manage competition within the lexicon. If this assumption is correct, individuals with low scores on this measure would be predicted to experience significant delays in producing a coherent utterance, which are likely to be reflected in low scores on the QPA speech rate measure and possibly also low scores on the struggle measure.

Third, difficulties with lexical selection, as indexed by low scores on the BLAST lexical selection measure, are also likely to introduce delays into the utterance planning process – which may be manifested in a reduced speech rate on the QPA, and may also be

associated with lower scores on the struggle measure, indicating greater difficulty forming a coherent utterance. In addition, since the lexical selection difficulty is likely to affect less frequent sentence elements, we would also predict closed class vocabulary to be disproportionately affected; consequently, scores on the closed class ratio should be abnormally *high*.

Table 3.1 summarises our predictions regarding the relationships between skills from the BLAST and measures obtained on our selected QPA measures.

Table 3.1.

Specific predictions regarding the expected relationships between the core skills contained in the BLAST and our selected QPA measures, as assessed using a Pearson's correlation

BLAST Core Skill	Quantitative Production Analysis (QPA) measure	Prediction
Verb Retrieval	Struggle Measure	A significant, positive correlation between verb retrieval and the struggle measure, as assessed using a Pearson's correlation
	Closed Class Ratio	A significant, positive correlation between verb retrieval and closed-class words
	Speech Rate (WPM)	A significant, positive correlation between verb retrieval and speech rate (poorer verb retrieval associated with lower speech rate)
Goal-driven Response Selection	Struggle Measure	A significant, positive correlation between goal-driven response selection and the struggle measure
	Speech rate	A significant, positive correlation between goal-driven response selection and speech rate (WPM)
Lexical Selection	Struggle Measure	A significant, positive correlation between lexical selection and the Struggle measure
	Speech rate (WPM)	A significant, positive correlation between lexical selection and speech rate (WPM)
	Closed Class Ratio	A significant, negative correlation between lexical selection and closed-class words

Turning now to the predictions with respect to comprehension – the TROG measures – our predictions were as follows. Our first set of predictions was that a reduced capacity to temporally store phonological information – as measured in the BLAST verbal STM skill measure - will result in deficits comprehending longer and/or more complex sentences. Consequently, low scores on the BLAST verbal STM measure are likely to be associated with low scores on the TROG "Total blocks passed" measure (a measure of overall accuracy). We would also predict patients to be disproportionately affected on complex items that have a high-retention load (below we describe how we define this).

Our second set of predictions was that a reduced capacity to identify auditory words – as measured on the BLAST auditory word recognition measure - will be associated with lower overall accuracy on the TROG total number of blocks passed measure, and also with difficulty on complex items that have a high retention load. However, in addition, we also predicted that those with low auditory word recognition scores would be affected across the board, even on items consisting of individual lexical elements (e.g., nouns, verbs) and grammatical elements (e.g. the cat the cow chased was black) – that is, a difficulty recognising auditory words was predicted to result in a greater frequency of lexical and grammatical errors.

Finally, turning to our BLAST measure of goal-driven response selection, we suggested above that difficulties with this skill – which involves managing competition between lexical elements planned for production – may also impact on comprehension, particularly when the sentences contain multiple lexical elements. Consequently, we predict that low scores on goal-driven response selection will be associated with a difficulty comprehending sentences with a number of lexical elements, resulting in lower scores on the total number of blocks passed. Again, sentences with a high retention load are also likely to be disproportionately affected, because these sentences tend to contain greater numbers of lexical elements, and may possibly also be more ambiguous as to the thematic role of each element.

Table 3.2 summarises our predictions regarding the relationships between skills from the BLAST and measures obtained on our selected TROG measures.

Table 3.2.

Specific predictions regarding the expected relationships between the core skills contained in the BLAST and our selected TROG measures

BLAST Core Skill	Test for the Reception of Grammar (TROG) measure	Operationalised prediction
Verbal short-term memory	Total Number of blocks passed	A significant, positive correlation between the total number of blocks passed and T scores on verbal short-term memory
	High retention load errors	A significant, negative correlation between verbal short-term memory and high-retention load errors
Auditory-word identification	Total Number of blocks passed	A significant, positive correlation between auditory-word identification and total number of blocks passed
	Retention Load errors	A significant negative correlation between auditory-word recognition and retention load errors (both low and high conditions)
	Grammatical and lexical errors	A significant negative correlation between auditory-word recognition and grammatical and lexical errors
Goal-Driven Response Selection	Total Number of blocks passed	A significant, positive correlation between goal-driven response selection and the total number of blocks passed
	High retention load errors	A significant, negative correlation between goal-driven response selection and high retention load errors

<center>**Method**</center>

The two key sentence processing tasks – the Quantitative Production Analysis (QPA) and the Test for the Reception of Grammar (TROG) – were administered to each participant during the long-term follow-up assessment. In total, 28 patients completed these tasks. The administration order of these tasks is outlined in Chapter two. These two tasks are described in detail below.

<center>**Quantitative Production Analysis (QPA)**</center>

Materials and Procedure. The Quantitative Production Analysis (Saffran, Berndt & Schwartz, 1989) involves individuals recounting a familiar fairy-tale, such as Cinderella. We administered the task according to the instructions set out in the administration manual by Saffran, Berndt and Schwartz (1989). To summarise, the aim of the task is to elicit at least 150 words of speech (not including repeats, filler words and fragments of words that were never completed). If the participant stops short of this target, the examiner can give prompts, such as, "*and what happened next*", or "*tell me about that*" (these prompts are excluded from the overall word count and speech rate measures). The participant is reminded that they are not being assessed on how accurately they recall the story. All narratives were recorded using Audacity software (Audacity Team, 2008) for later transcription.

Response Scoring and Analysis. All response scoring and analyses were performed according to the QPA scoring protocols (Saffran et al., 1989), whilst the control data was obtained from Rochon, Saffran, Berndt, and Schwartz (2000). We selected the following three measures to assess the quantitative aspects of speech production, which are further outlined in Table 3.3: i) *rate of speech;* ii) *the proportion of closed class words*; and iii) *the struggle measure* (the amount of difficulty eliciting a connected and coherent narrative). The rate of speech is calculated from the entire speech sample and is simply the total number of words (excluding examiner prompts and fillers, e.g. *ahhh*, *er*), divided by the total time to elicit the narrative (in minutes). The proportion of closed class words was calculated from the first (150 +/- 10) "narrative words", which further excludes words that do not contribute to the narrative, such as *repetitions*, *neologisms, habitual starters, conjunctions*, and words that are later corrected. Words in this narrative are classed as closed class if they were not function words - that is nouns, verbs, adjectives and adverbs were excluded. Further, elements such as *'be', 'do',* and *'have'* were considered open-class words when they appear as the main verb, and were therefore not included in the overall closed class count (Saffran et

<center>128</center>

al., 1989). The *struggle measure* expresses the proportion of narrative words produced by the speaker (150 +/- 10) as a proportion of the total number of words.

Each individual's score on each measure was compared to that of the appropriate normative group using Crawford and Howell's (1998) modified t-test. As discussed previously, this method was selected as it treats the control sample as statistics rather than parameters, and has been confirmed as a robust method that controls for skewedness in the data, as well as controlling for type 1 error rate, regardless of the size of the control sample (see Crawford & Garthwaite, 2006).

Table 3.3.

Each of our selected measures and definitions for the QPA task

Quantitative Production Analysis (QPA)	Definition
Speech rate	Total number of words/ total time to produce narrative (in minutes)
Proportion of closed-class words	Proportion of narrative words that were closed class (number of closed-class words/ number of narrative words)
Struggle measure	Total number narrative words divided by the proportion of total words uttered

Test for the Reception of Grammar

Materials. The Test for the Reception of Grammar (TROG; Bishop, 1989; 2003) is administered orally according to the original protocols originally outlined by Bishop (1989). The task is divided into two sections; a comprehension section, and an experimental section. The comprehension section consists of 48 pictorial items that are divided into six blocks, each of which assess three major classes of lexical elements: nouns, verbs, and adjectives. The ability to understand these elements is considered necessary for the comprehension of more complex utterances, therefore this section provides a baseline measure of this capacity. In contrast, the experimental condition examines the comprehension of complete sentences. Each item consists of four pictures, one which depicts the target sentence, and three that

depict sentences that differ from the target along some important dimension. In some instances, the distractor pictures depict contrasting sentences that contain different lexical elements, and in other instances, they depict contrasting sentences that contain the same elements within a different grammatical structure (see Table 3.4). There are a total of 80 items, which are organised into blocks of four based on the type of contrast they examine. The blocks are arranged in order of increasing difficulty. Table 3.4 presents a summary of the various types of items in the experimental section of the TROG test.

Table 3.4.

Each of the types of experimental contrasts that are measured in the TROG, including whether the contrasts contain lexical (L) or grammatical (G) elements on the relevant blocks (- denotes that the block did contain lexical or grammatical distractors or elements). Also shown is the target element/sentence and examples of the distractor element/sentence. Finally, high and low-retention loads are also shown (- denote that the contrast was not classed as having either a high or low retention load).

Block/ type of contrast	Target Element/Sentence	Example Distractor (lexical)	Example Distractor (grammatical)	High or Low Retention load
Nouns	"*shoe*"	tree	-	-
Verbs	"*eating*"	kicking	-	-
Adjective	"*long*"	short	-	-
Two elements	"the boy is running"	the cat is <u>sitting</u> (L)	-	-
Negative	"the boy is not running"	the horse is not running (L)	the boy is running (G)	-
Three elements	"the boy is jumping over the box"	the <u>horse</u> is jumping over the box/ the boy is	-	-

		jumping over the flowers (L)		
Plural/ singular pronoun	"they are sitting on the table"	they are sitting on the mat (L)	he is sitting on the table (G)	-
Reversible active	"the girl is pushing the horse"	the girl is pushing the man (L)	the horse is pushing the girl (G)	-
Masculine/ feminine pronoun	"She is sitting on the chair"	she is sitting on the grass (L)	he is sitting on the chair (G)	-
Singular/ plural noun	"The cats look at the ball"	the people look at the ball (L)	the cat looks at the ball (G)	-
Comparative	"The knife is longer than the pencil"	the flower is longer than the pencil (L)	the knife is the same size as the pencil (G)	-
Reversible passive	"The girl is chased by the horse"	the girl rides the horse (L)	the girl chases the horse (G)	Low
In and on	"The cup is in the box"	the cup is on the box	-	Low
Post-modified subject	"The boy chasing the horse is fat"	the boy being chased by the horse is fat	-	High
X but not Y	"The box but not the chair is red"	the chair but not the box is red	-	High
Above and below	"the pencil is above the flower"	the pencil is below the flower	-	Low
Not only X but also Y	"Not only the bird but also the flower is blue"	only the flower is white/ not only the bird but also the flower is white	-	High
Relative clause	"The pencil is on the book that is yellow"	the pencil is under the book that is yellow	-	High

Neither X nor Y	"Neither the dog nor the ball is brown"	<u>only</u> the dog is brown	-	High
Embedded sentence	"The book the pencil is on is red"	the book the pencil is <u>under</u> is red	-	High

Table 3.5 summarises the measures obtained from the TROG tests. These measures include two measures that we created for the purposes of this study: High retention load and low retention load errors. These measures were created as they allowed for the measurement and comparison of performance on more complex items that place greater demands on the retention of relatively more difficult sentences compared to those that make comparatively fewer demands. To obtain an overall total for each condition, the total number of errors for each corresponding block was calculated. We then calculated a difference score, by subtracting the number of errors on the high retention load condition to the number of errors on the low retention load condition.

Table 3.5.

Each of our selected measures and definitions from the TROG.

Test for the Reception of Grammar	Definition
Total number of blocks passed	Total number of blocks successfully passed. A block is passed if all four items in that block are responded to correctly.
Proportion lexical confusions	Proportion of responses (number of lexical errors divided by the number of blocks completed) where the sentence selected differed in lexical content from the target (e.g., *The <u>boy</u> is chasing the sheep - > The <u>dog</u> is chasing the sheep*).
Proportion grammatical confusions	Proportion of responses (number of grammatical errors divided by the number of blocks completed) where the sentence selected differed with respect to the grammatical relationships amongst elements (e.g., *The boy is chasing the sheep - > The sheep is chasing the boy*).

Total high retention load errors	Total errors on items classified as having a high retention load, defined as those that were more complex and placed greater demands on comprehension (see Table 3.4 for examples).
Total low retention load errors	Total errors on items classified as having a low retention load, defined as those that were comparatively less complex in nature and placed comparatively less demands on comprehension (see Table 3.4 for examples).
Difference high and low retention load	The difference in error rates between high retention load conditions and low retention load conditions.

Procedure. The TROG was administered according to the standard procedure set out in the test manual (Bishop, 1989). In the comprehension section, participants were first asked to name each picture. For example, on the noun and adjective contrasts, participants were asked "what are these things?" (e.g. *chair, big)*. On the verb contrasts, participants were asked "can you tell me what they are doing" (e.g. *running*, *skipping*). For any items that were incorrect, participants were then presented with a spoken phrase and asked to point to the corresponding picture. For example, on the noun contrasts, they were instructed to "show me the… *shoe*", whilst on the verb contrasts, they are asked, "which one is… *skipping*?". Errors during this stage are corrected.

Following successful completion of the comprehension section, the experimental condition was administered. In this condition, the examiner read aloud a phrase or sentence, and participants were asked to point to the corresponding picture from an array of four images (Figure 3.1). On this condition, testing was discontinued if five consecutive blocks were failed.
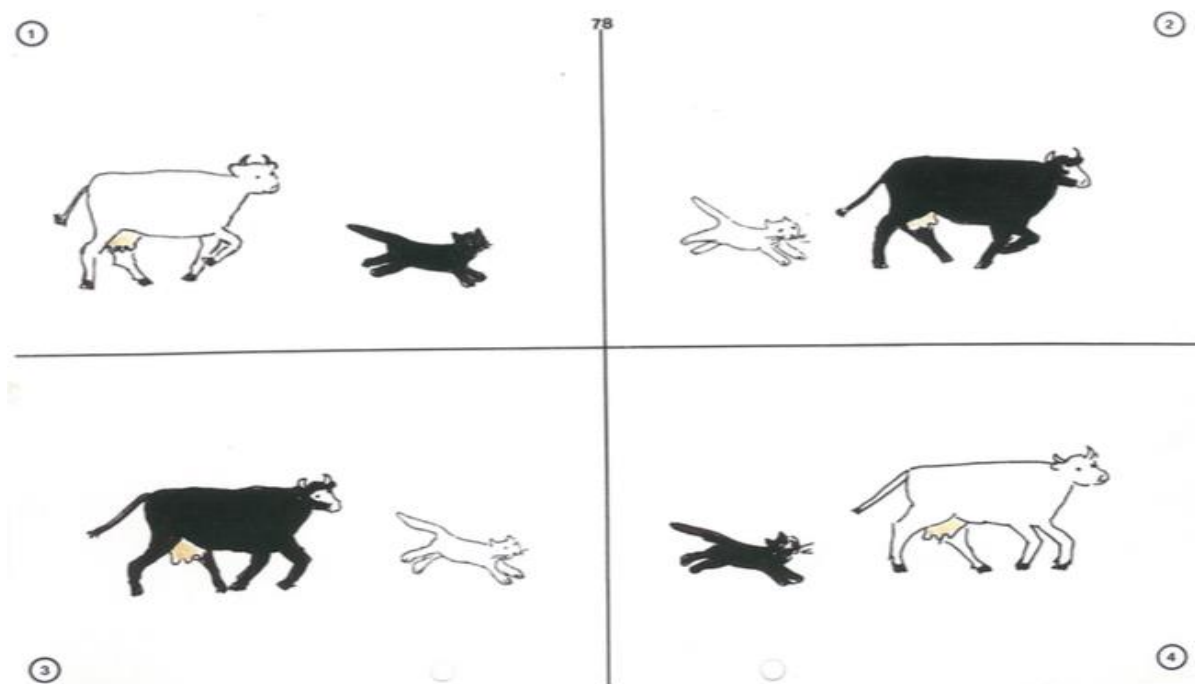
Figure 3.1. An example of a high-retention load condition in the experimental condition. In this example, the examiner reads aloud the sentence "*the cat the cow chases is black*", and participants are required to point to the corresponding picture. In this case, the correct answer would be image one.

Response scoring. In the comprehension section, a response was scored as correct if the participant identified the appropriate noun, verb or adjective associated with the pictorial image, and was able to produce its name. In the experimental section, a response was scored correct if the appropriate picture was selected. Responses that were changed from incorrect to correct, or that needed to be repeated by the examiner were considered correct. Scores were obtained for the measures listed in Table 3.5.

Results

Sentence-Level Speech Production: Overall Findings

Considering first the sentence-level production measures, Table 3.6 shows each participant's mean score on each of the key QPA measures. Scores accompanied by an asterisk indicate that value was significantly different from those for the corresponding healthy control group, using Crawford and Howell's (1998) modified t-test. It can be seen that, on measures of *speech rate (WPM)*, and *proportion of closed-class words*, no patients showed a significant impairment relative to controls. However, one patient was significantly impaired on the *struggle measure* (S.O., left parietal glioblastoma). A breakdown of S.O.'s results shows that 34.6% of words were classed as "non-narrative" words, and included repetitions (e.g., "*there's,. there's a big..*"), and self-repairs (e.g. "*the girls… three sisters..*"). In total, 59% of S.O.'s narrative consisted of closed-classed words, whilst 41% were open-class words. Figure 3.2 displays a short sample of S.O.'s speech; as is evident, S.O.'s narrative displayed many qualitative impairments that were not detected in our quantitative measures.

Table 3.6.

Individual scores for each of the QPA and TROG measures respectively. N.B Higher struggle scores reflect proportionately more narrative speech, whilst lower struggle scores reflect less narrative speech, and thus are indicative of a greater struggle. (N.B. LGG = low-grade glioma, HGG= High-grade glioma).

Patients	Tumour Location	Tumour Type	WPM	Closed-Class Words	Struggle Measure	Total Blocks Passed	Proportion Grammatical Errors	Proportion Lexical Errors	Difference high and low
			QPA			**TROG**			
B.P	L Frontal	Meningioma	164	0.62	0.85	20	0	0	0
R F	L Frontal	Meningioma	160	0.70	0.95	20	0	0	0
T.D	L Frontal	LGG	131	0.59	0.9	20	0	0	0
L.C	L Frontal	Astrocytoma	122	0.58	0.84	19	0	0	1
DP	L Frontal	Astrocytoma	154	0.60	0.95	17	5	0	0
Average			*146*	*0.62*	*0.90*	*19.2*	*1*	*0*	*0.2*
N.H	R Frontal	Meningioma	113.5	0.58	0.88	19	0	5	0
D.F	R Frontal	LGG	141.1	0.64	0.91	18	0	0	2
BS	R Frontal	Meningioma	134.5	0.61	0.86	19	0	0	1

L.W	R Frontal	Meningioma	134.1	0.62	0.88	19	0	5	0
G.P	R Frontal	Meningioma	109.1	0.60	0.9	6	6.25	43.75	-1
R.S	R Frontal	Meningioma	162	0.65	0.85	20	0	0	0
WR	R Frontal	Meningioma	159	0.55	0.95	16	5	0	0
S.N	R Frontal	Meningioma	194	0.58	0.96	19	0	5	-2
Average			*136*	*0.60*	*0.88*	*17*	*1.41*	*7.34*	*0.375*
K.W	L Temporal	Meningioma	141.5	0.57	0.78	20	0	0	0
S.H	L Temporal	Glioblastoma	174.8	0.66	0.75	15	5	15	1
G.M	L Temporal	Glioblastoma	156.4	0.67	0.87	17	0	5	3
C.A	L Temporal	Meningioma	106.8	0.60	0.88	17	0	0	3
Average			*144*	*0.62*	*0.81*	*17.25*	*1.25*	*5*	*1.75*
B.D	L Parietal	Meningioma	104.9	0.58	0.91	16	0	5	3
C.R	L Parietal	Meningioma	144	0.60	0.99	19	5	0	0
A.V.G	L Parietal	Meningioma	151.3	0.55	0.86	19	0	0	1
N.O.H	L Parietal	HGG	152.8	0.61	0.84	17	5	0	2

S.O	L Parietal	Glioblastoma	182.8	0.59	*0.65**	18	0	5	1
K.G	L Parietal	LGG	184	0.64	0.91	18	5	0	1
Average			*153.3*	*0.59*	*0.59*	*17.25*	*2.5*	*1.7*	*1.3*
S.G	R Posterior	Astrocytoma	108.5	0.54	0.76	18	0	0	2
A.E.K	R Posterior	HGG	74.3	0.60	0.96	19	0	0	1
J.B	R Posterior	Meningioma	134.7	0.65	0.94	19	0	0	1
P.C	R Posterior	Meningioma	132.5	0.55	0.82	20	0	0	0
Average			*112.5*	*0.58*	*0.868*	*19*	*0*	*0*	*1*
AE	R & L Frontal	Meningioma	167.61	0.71	0.85	17	0	0	4

*p<.05.

> *"Ok, we'll work on Cinderella where..// where the girls were going to go … // three sisters were going to go to … ahh the ball, I think it was two beautiful looking sisters and one ugly looking sister and um she got..// ah went out to the ball and had some spell and um managed to get a beautiful dress.. beautiful um ch // chariot and changed in from a um… into a pumpkin or something and they went to this ball and um.. danced with a prince who f/ ff/… fell in love with her and she dropped a slipper on the way out and um yeah there's..// there's a big..// big ahh hunt on to track this..// who this person was and it was the ugly sister or the three sisters who they thought would never go to the ball and um.. ah… so I think she finally got tracked down with the loose.//. ah lost slipper and um..ahh was deemed to be the*

Figure 3.2. A short sample of S.O.'s speech when he was asked to produce the story of Cinderella.

Table 3.7 presents the mean scores and standard deviations for our selected QPA measures across the following five localisation groups, as defined in Chapter 2: left frontal (n= 5), right frontal, (n= 8) left temporal (n= 4), left parietal (n= 6), and right posterior (n= 4)[14]. Using SPSS Software, an omnibus one-way Analysis of Variance (ANOVA) revealed no significant effects of localisation group for any of our selected QPA measures. It should be borne in mind that the number of patients in some groups was small, which may have limited statistical power to detect significant effects.

---

[14] One patient, AE, was excluded from this analysis due to the presence of multiple lesions

Table 3.7.

Average scores and standard deviations (in parentheses) across the five anatomical groups.

	Left Frontal	Right Frontal	Left Temporal	Left Parietal	Right Posterior
Number per group	5	8	4	6	4
Speech rate (WPM)	146.3 (18.7)	143.5 (27.8)	144.9 (28.8)	153.3 (29.4)	112.1 (28.1)
Proportion of Closed Class words	0.38 (.050)	0.39 (.027)	0.37 (.050)	0.40 (.031)	0.41 (.051)
Struggle Measure	0.90 (.053)	0.86 (.039)	0.81 (.065)	0.85 (.113)	0.86 (.092)

Table 3.8 shows the average scores and standard deviation for each of our QPA measures as a function of tumour grade. Using the methods outlined in Chapter 2, patients were grouped into one of the following two broad groups, based on the histological grading of their tumour: high-grade (n= 5) and low-grade (n= 23). Using SPSS Software, an independent samples t test revealed no significant differences between the two groups on any of our selected QPA measures.

Table 3.8.

Average scores and standard deviations (in parentheses) as a function of tumour grade.

	Low-Grade Tumours	High-Grade Tumours
WPM	141.54 (24.59)	148.22 (43.17)
Proportion of Closed-Class Words	0.60 (.05)	0.63 (.4)
Struggle Measure	0.89 (.06)	0.81 (.12)

In summary, there was little evidence of impairment in our groups on the key QPA measures. There are two possible explanations for this finding. First, it is possible that the (probably mild) deficits measured in the various BLAST single word assessments are not

sufficiently severe to impact on everyday speech. Second, it is possible that some or all of these individuals *are* limited in their everyday speech, but that the variables measured in the QPA may be insufficiently sensitive at detecting those subtle deficits. Indeed, the QPA is a relatively blunt instrument when it comes to assessing speech quality. Specifically, the scoring protocol counts only the rates of certain word types, but it does not assess accuracy, nor appropriate or grammatical correctness of those words.

Relationship to Core Cognitive Language Skills

We next tested our specific predictions concerning the relationship between selected QPA measures and individual BLAST core skill measures. Table 3.9. displays the uncorrected Pearson's r values for the correlations that address each of our predictions. Consistent with our original prediction, a significant, negative correlation was observed between the BLAST *lexical selection* measure and the QPA *closed class measure,* indicating that poor lexical selection is associated with a reduced ability to produce rich connected speech containing a range of open class, content words.

Table 3.9.

Pearson's correlation coefficients for the predicted relationships between the BLAST skills and selected QPA measures

		Speech Rate (WPM)	Proportion Closed Class Words	Struggle Measure
Lexical Selection	Correlation	-0.19	-0.420*	0.452*
	P Value	*.34*	*.026*	*.016*
Verb Retrieval	Correlation	0.08	0.126	0.399*
	P Value	*.7*	*.522*	*.036*
Goal-Driven Response	Correlation	-0.06	-0.210	0.129
	P Value	*.8*	*.292*	*.52*

* p<.05

Again, consistent with our original prediction, a significant, positive correlation was observed between our BLAST *lexical selection* measure and the QPA *struggle measure.* This finding indicates that poor lexical selection was associated with greater evidence of struggle in connected speech, as evidenced by frequent false starts, repairs and word repetitions. Further, consistent with our predictions, a significant positive correlation was observed between the BLAST *verb retrieval* measure and the QPA *struggle measure*, indicating that the limitations in the capacity to retrieve verbs also result in greater evidence of struggle in connected speech. However, contrary to our original hypotheses, no other significant associations were observed. Figure 3.3. displays the relevant scatterplots for each of the significant correlations. Regression analyses were also conducted, as evident in the Figure. As is evident from the Figures, these significant correlations did not appear to be driven by just one patient.
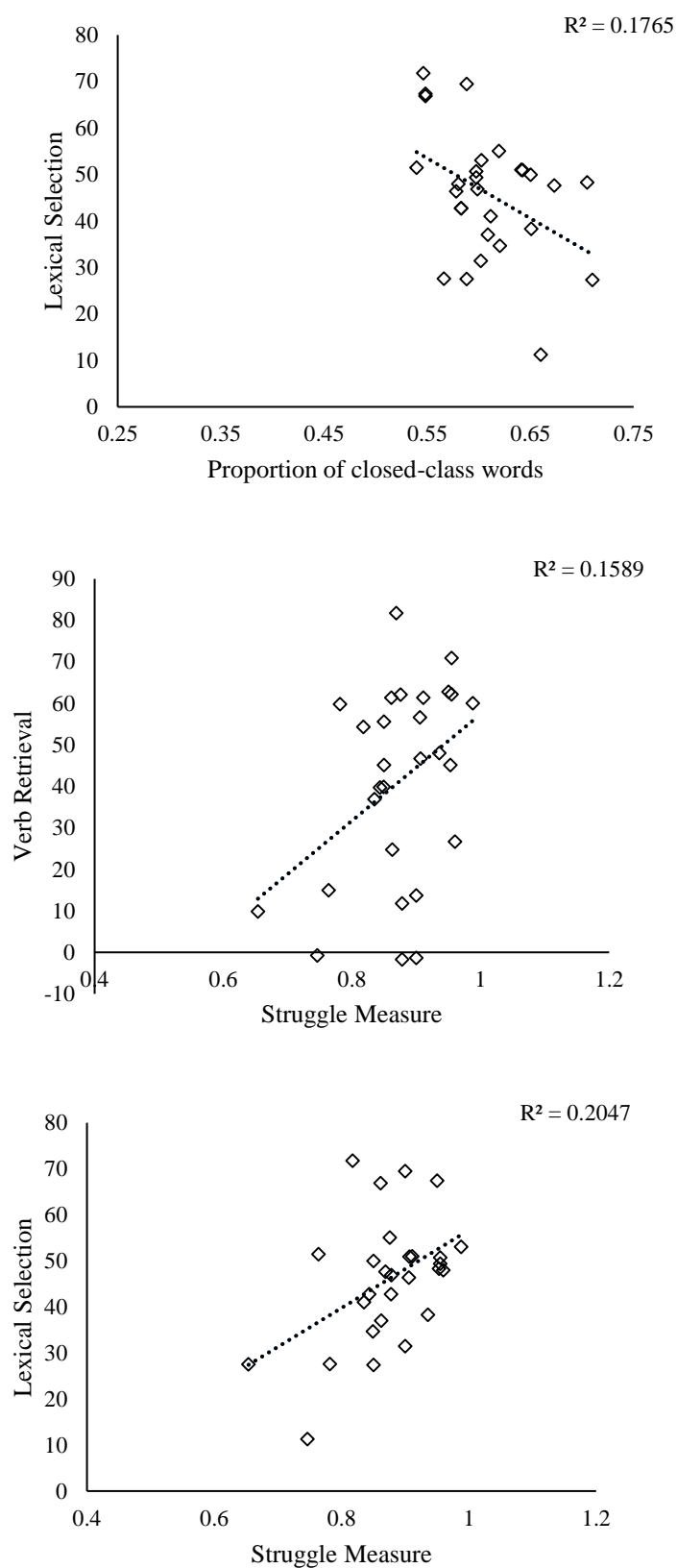
Figure 3.3. Scatterplots of the significant correlations and regression analyses between the core skills and selected QPA measures. The x axis represents the QPA measures whilst the y axis represents the BLAST core skills measures.

143

Sentence-Level Comprehension: Overall Findings

Table 3.6. presents patients' scores on the key measures from the TROG test. As is evident, only one patient (G.P: right frontal meningioma) showed greater impairment on any measure – specifically, G.P. failed more TROG blocks than the remaining patients, with his errors including: one grammatical error; seven lexical errors; two errors on constructs measuring high retention load; and three errors on constructs measuring low retention load. Importantly, G.P. did not appear to be disproportionately impaired on items with a high retention load. Taken together, these findings suggest that G.P.'s difficulty is most likely at the lexical level.

Using the methods described above, an omnibus one-way ANOVA revealed no significant effects of tumour localisation (left frontal, right frontal, left posterior, left temporal, and right posterior) on any of our selected TROG measures. Moreover, an independent samples t test also revealed no significant effects of tumour grade (high versus low-grade) on our TROG measures.

Relationship to core cognitive skills

Turning now to the main objective of this section, which was to investigate the relationship between each of our selected TROG measures and the BLAST core skills measures. To do this, we performed a Pearson's correlation using SPSS Software. Table 3.10. displays the uncorrected Pearson's r values for the correlations that address our predictions. The Table also presents results for correlations that were approaching significance based on uncorrected p values ($p < .05$).

Table 3.10.

Pearson's correlation coefficients for the predicted and significant relationships between the BLAST skills and selected TROG measures

		Total Blocks Passed	Proportion Grammatical Errors	Proportion Lexical Errors	Difference High vs Low Retention
Auditory Word Recognition	Correlation	0.087	-0.104	0.148	-0.347
	P Value	*.66*	*.6*	*.453*	*.070*
Lexical Selection	Correlation			-0.372	
	P Value			*.051*	
Verbal SMT	Correlation	-0.155			
	P Value	*.439*			
Verb Retrieval	Correlation			-0.457*	
	P Value			*.015*	
Goal-Driven Response	Correlation	-0.032			
	P Value	*.874*			

* p<.05

As seen in Table 3.10., there were few significant associations between the BLAST core skill measures and our selected measures on the TROG test. The only significant association was observed between *verb retrieval* and the *proportion of lexical errors,* wherein a significant negative correlation was observed. That is, better verb retrieval skills were associated with fewer confusions between sentences differing in their lexical content. This relationship was not originally predicted. However, as the scatterplot in Figure 3.4. shows, this correlation appears to have been driven by just a single patient, who presented with a right frontal meningioma: G.P.
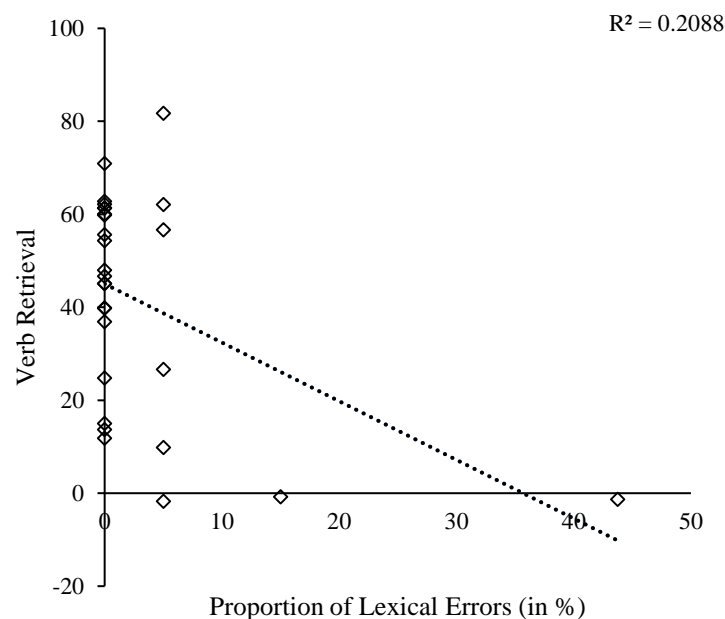
Figure 3.4. Significant negative correlation between Verb Retrieval and the proportion of lexical errors. The x axis represents the TROG sentence-comprehension measure, whilst the y axis represents the BLAST verb retrieval measure.

As seen in Table 3.10., the correlation between *lexical selection* and the *proportion of lexical errors* just failed to reach significance ($p = .051$). This finding, if confirmed in subsequent studies, would be particularly interesting given the measure of lexical selection in the BLAST is conceptualised as involved in language production only, and may suggest that similar skills are recruited during certain types of comprehension tasks. Similarly, the correlation between *auditory word recognition* and *retention load difference score* also just failed to reach significance ($p = .070$). Again, this finding, if confirmed in subsequent studies, would suggest that difficulties with auditory word recognition become exaggerated in

situations where multiple words must be processed and retained within close succession. Inconsistent with our original predictions, no other significant findings were observed. Possible reasons for this are discussed below.

Discussion

The overall objective for this Chapter was to investigate the relationship between connected, sentence-level speech, and more elemental core skills, such as those tested in the BLAST. We also explored general patterns of impairment on each of our connected speech measures. Overall, the incidence of significant impairment was low on both the sentence-level production and comprehension measures - the QPA and TROG respectively. There were also no reliable effects of tumour localisation or tumour grade on any of the selected measures.

These findings were surprising, given that some patients continued to show long-term impairment on the BLAST core skills, suggesting unresolved long-term language deficits. They are also somewhat at odds with previous studies that have reported impaired propositional speech in neurological tumour populations, even in the context intact naming, and repetition (e.g., Satoer et al., 2014).

One possible explanation for the above results is that our selected sentence-level measures were insufficiently sensitive. For example, the QPA emphasises the rate of speech and the content of spontaneous speech in terms of the broad grammatical classes represented. Narratives do not need to be grammatically correct, nor lexically or semantically appropriate to receive scores within the normal range. This was illustrated in one patient, S.O. who presented with a left parietal glioblastoma. S.O. exhibited fragmented speech that contained few complete sentences, yet, he still scored within the normal range on key QPA measures. Accordingly, despite the QPA being a well-validated measure of sentence level language production in stroke aphasia populations – particularly those with non-fluent aphasia - the subtle language impairments that are evident in neurological tumour populations may not be adequately detected using QPA measures.

Further, given many of our patients exhibited ceiling effects on the selected TROG measures, it is possible that this simple picture choice task was not sufficiently challenging for our participants. Accordingly, future studies of sentence-level comprehension should endeavour to include a measure that is arguably more challenging, either with respect to the speech sample to be analysed (e.g., paragraph, short stories), or with respect to the way in which comprehension is assessed (e.g., using event-related potentials to identify specific "markers" associated with correct semantic or syntactic integration).

However, another possibility is that there was insufficient genuine impairment in our sample at the time of testing (long-term follow-up). For example, most patients exhibited normal long-term follow-up scores on all but one or two of the BLAST core skill measures, particularly those hypothesised to be most involved in sentence-level language. Therefore, it is also possible that these deficits were insufficiently severe or numerous to significantly impact on the connected speech level. Finally, it should be borne in mind when assessing the results of this study that the sample size was small, which limits the statistical power to detect significant effects.

Nevertheless, there were some interesting phenomena that are worth commenting upon here. First, considering the QPA speech production measures, one patient who received a particularly low score on our QPA speech rate measure was A.E.K., who had a high-grade, right posterior malignant glioma. Although A.E.K.'s lesion impinged on the right motor strip, and she presented preoperatively with a left-sided facial droop, these motor deficits are unlikely to fully explain her slowed spontaneous speech (for example, she performed normally on the articulatory agility task). Her reduced speech rate is more likely to reflect other, non-linguistic factors. For example, A.E.K. demonstrated generally slowed processing on a number of cognitive tasks, including those not involving speech, and this may have been related to her recent radiation therapy (see Chapter 6). Moreover, Alexander (2006) suggests that the production of an extended discourse requires the speaker to be able to sustain an overall schema of the task and the message to be communicated throughout; consequently, if a person's processing is generally slowed, the schema may decay before the narrative is complete. This is likely to result in frequent hesitations and pauses at shift points in the narrative (Alexander, 2006). If this is indeed contributing to A.E.K.'s low speech rate, then it may suggest that difficulties in the activation of non-verbal responses result in generalised difficulties affecting the rate of spoken speech.

Turning now to the TROG sentence comprehension measures, only one patient, G.P., showed disproportionately lower scores relative to the group. G.P., who presented with a large right frontal meningioma, scored poorly on the *total number of blocks passed* relative to the other patients. Importantly however, he was not significantly impaired on our core skill measures of *verbal short-term memory* or *auditory word identification,* both of which were predicted to correlate with the total number of blocks passed. A more likely explanation for G.P.'s difficulties on the TROG is that he had difficulty following the task instructions and maintaining the task set. Specifically, G.P. exhibited impairments on a range of cognitively-

demanding BLAST tasks, including the Stroop task and its non-verbal analogue of the anti-saccade task (Chapters 4 and 5). Indeed, it has been argued that picture matching tasks, such as that utilised in the TROG, place quite considerable demands on cognitive control functions, because the participant must not only be able to understand the sentence itself, but must also select its best matched picture, while at the same time inhibiting the potential distracting effect of the other picture options. In the following Chapters, we examine G.P.'s general cognitive control skills in greater detail. The next section considers the specific outcomes of this analysis with respect to the key predictions.

Relationship between sentence-level measures and BLAST skill scores.

We made several predictions regarding the relationship between BLAST core skill scores and selected QPA measures. Our first prediction, that there would be a significant relationship between BLAST *verb retrieval* scores and the QPA *struggle measure*, was supported. This finding suggests that the capacity to retrieve verbs from the lexicon is crucial for effective, fluent and concise speech. These results are consistent with sentence production theories that place particular emphasis on the role of verbs at both the conceptual and grammatical levels of sentence structure (Marshall et al., 1998; see also Ahrens, 2003; Berndt et al., 1997; Levelt 1989; 1999; Pickering & Branigan, 1998; Shapiro & Levine, 1990; Trueswell & Kim, 1998). Accordingly, the significant correlation between verb retrieval and the QPA struggle measure supports previous suggestions that verbs do indeed act to specify the relationship between nouns and their thematic role in order to formulate a coherent argument structure (Marshall et al., 1998; Rofes and Miceli 2014). Consequently, at the elemental level, difficulty selecting and retrieving verbs is likely to underpin the formulation and construction of sentence-level speech. Such findings suggest that this process can be decomposed to the core skills level. Accordingly, interventions which target the individual verb level specifically may indeed be particularly useful at ameliorating sentence-level expressive difficulties, at least in those with evidence of difficulty at this level.

Our second and third predictions were that there would be a significant relationship between BLAST lexical selection scores and two QPA measures: the proportion of *closed-class words* and the *struggle measure*. These two predictions were also supported. This suggests that poor lexical selection skills manifest themselves in connected speech in two ways. First, they result in speech that is generally lacking in open class, content words, and second, they are associated with a higher abundance of false starts, repairs and word repetitions. This will in turn impact on the overall informativeness of the narrative. From a

clinical perspective, these findings also suggest that rehabilitation targeted at the individual lexical level may be particularly useful at generalising to the sentence-level in affected individuals.

However, our fourth prediction, that there would be a significant relationship between *lexical selection* and the following QPA measures was not supported: *words per minute* and the *proportion of closed-class words*. The process of lexical selection occurs relatively quickly at a rate of between 1-3 words per second (Butterworth, 1989; 1992; Levelt, 1989); consequently, a difficulty selecting and retrieving lexical items would be expected to result in reduced speech output. Of course, we need to bear in mind that the QPA is relatively unconstrained, and it is possible to produce acceptable speech on the task by relying on relatively high-frequency words, semantically empty verbs (e.g. have, make, do), and even semantically empty nouns, such as "thing" and "type". Interestingly, S.H., who scored particularly poorly on the BLAST lexical selection measure, was able to maintain a high speech rate on the QPA, although most of the open class words in his output were very high-frequency words. A similar pattern was evidenced in other patients who scored poorly on the BLAST lexical selection measure. To overcome this issue, future studies should endeavour to include a more intermediate task of sentence-level speech production that is relatively more constrained, yet allows for the scoring of informative accuracy and grammatical well-formedness (Wilshire, Lukkien, & Burmester, 2014). These types of tasks include structured picture event tests and the Sentence Production Test (SPT: Wilshire et al., 2014); this may allow for more systematic comparison between lexical selection and sentence-level production.

Our fifth prediction, that there would be a significant association between BLAST *goal-driven response selection* scores and the QPA *struggle measure,* was not supported. We predicted a relationship between these two measures based on the notion that a dedicated control system is called upon to manage competition within the lexicon when multiple elements are activated (Biegler et al., 2008; Robinson et al., 2005). Of course, we need to bear in mind that the BLAST goal-driven selection measure is defined relatively broadly, and might itself decompose into several cognitive elements (e.g., response inhibition, vs. effortful, controlled lexical selection; see, e.g., Nigg, 2000). It may not be possible to generate more specific predictions regarding the role of goal-driven response selection in naturalistic speech until we know more about the nature of the processes underlying this capacity.

Turning now to the TROG measures, a significant negative correlation was observed between *verb retrieval* and the *total number of lexical errors* on the TROG task*, suggesting that better verb retrieval skills are associated with fewer confusions between sentences differing in their lexical content. This correlation was not originally predicted, however a scatterplot that displays this data showed that this effect was largely driven by one single individual, G.P., who scored particularly poorly on both measures.

None of the other predictions listed in Table 3.2 were supported. For example, we did not find evidence for an association between our TROG measures and the BLAST *auditory-word identification* measure. This finding is perhaps not surprising, given that none of our participants scored below the normal range on the relevant BLAST measure, and only one participant (G.P.) showed poor scores on the corresponding TROG measure. It is also worth bearing in mind that the auditory word recognition demands imposed by the TROG are likely to be low, and therefore, this measure is unlikely to be highly sensitive to selective impairments in auditory word recognition. Specifically, the BLAST core skill of auditory-word identification was measured through errors on the repetition task and phonological distractor items on the Picture-word Verification task, both of which are arguably phonologically driven (e.g. chair -> *cheque*). In contrast, items on the TROG manipulate grammatical and semantic aspects of a sentence using high frequency lexical elements (for example, "the boy jumped *over* the box") that rarely require the listener to make fine-grained phonological discriminations. Accordingly, future studies that aim to examine the impact of single word recognition deficits on sentence-level comprehension need to use a wider range of lexical elements and include both phonological as well as semantic distracters.

Moreover, contrary to our prediction, no significant correlation was observed between the TROG primary measure (number of blocks passed) and the BLAST goal-driven response selection measure. Again, the primary problem here is the near-ceiling performance of our participants on the TROG measures. Previous studies have posited an association between the resolution of temporary ambiguity and conflict in the lexicon (see Novick et al., 2009). However, in contrast to those temporarily ambiguous sentences used in studies such as Novick et al (2009; e.g. "*put the apple on the napkin into the box*"), the ambiguous sentences included in the TROG place relatively low demands on ambiguity resolution. In addition, and as we have argued above, the goal-driven response selection measure might be comprised of a number of distinctly different cognitive skills. Finally, it is also worth noting that G.P., the patient who scored poorly on a number of the TROG measures, did not score poorly on the

BLAST goal-driven response measure. This is surprising, given the interpretation we presented above for G.P.'s deficit, which was that it might reflect a difficulty following task instructions.

Finally, our prediction, that there would be a significant relationship between selected TROG measures and the BLAST *verbal short-term memory* measure, was also not supported. Again, a major problem here was the near-ceiling performance of patients on the TROG across the board. However, it is also worth considering this finding in the light of those reported by Gvion and Friedmann (2012), who investigated the relationship between phonological short-term memory and sentence comprehension in a sample of 12 individuals with conduction aphasia. Specifically, comprehension capacities were evaluated through relative clauses using sentence-picture matching and plausibility judgment tasks whilst phonological short-term memory was assessed using a rhyme judgment sentence test and a paraphrasing task that manipulated the ambiguity of words. Overall, despite considerable impairment on phonological short-term memory, the patients only exhibited semantic comprehension impairments when sentences required phonological reactivation; that is, during ambiguous sentences, all semantic meanings are activated. However, during the incremental processing of an ambiguous sentence, one meaning receives the greatest activation, whilst the others decay; therefore, in these types of ambiguous sentences, the initial semantic form cannot be used, and instead the phonological word form of the original word must be reactivated in order to re-access all possible meanings to comprehend the sentence (see Gvion and Friedmann, 2012). Such findings led Gvion and Friedmann (2012) to conclude that phonological working memory is only involved in comprehension under very specific circumstances that require phonological reactivation. These circumstances were not present in the relevant TROG items.

Conclusion. Overall, the results from the current investigation suggest that measures of core skills such as those assessed in the BLAST can indeed be predictive of performance on more naturalistic, connected speech tasks, but that their effects are only evident on certain types of tasks and measures. Scores on specific BLAST measures were predictive of several of the measures of connected speech derived from the QPA. Specifically, the significant correlation between BLAST verb retrieval and the QPA struggle measure suggest that at the elemental level, difficulty selecting and retrieving verbs is likely to underpin the formulation and construction of sentence-level speech. Further, the significant correlation between BLAST lexical selection scores and two QPA measures - the proportion of closed-class

words and the struggle measure - suggests that poor lexical selection will in turn impact on the overall informativeness of the narrative. However, these observations do not extend to sentence-level comprehension, at least not as it is measured using the TROG. It is always difficult to assess comprehension, because the results one obtains can be strongly influenced by the constraints of the task itself (e.g., the nature and number of distractors and kind of choice rule required). Accordingly, it is possible that the conditions of the TROG do not adequately reflect the normal language processes that occur during conversational settings; these limitations, as well as potential adjustments for future research, are further discussed in Chapter 6.

Chapter 4: Characterising Cognitive Control

Introduction

So far, the emphasis of the previous chapters has focused on language specifically. However, to fully understand the mechanisms that underpin complex language operations, it is also necessary to investigate higher-level, non-linguistic functions, such as cognitive control. Cognitive control is a general term used to refer to a collection of cognitive skills that are necessary to internally guide goal-oriented behaviours (Badre, 2008; Badre et al., 2009). As discussed in Chapter 1, theoretical models of cognitive control fall under different categories depending on the demands they emphasise (e.g. unitary, hierarchical, regional specialisation accounts). One such model emphasises the specific processing specialisations of the various prefrontal cortex (PFC) subregions. This account, proposed by Stuss and colleagues (1995; 2005), arguably provides a more comprehensive model of regional specialisation. Specifically, it draws upon the basic framework of Norman and Shallice's (1986) Supervisory Attentional System by proposing that specific subregions within the PFC subserve functionally distinct components of cognitive and behavioural control (see Chapter 1). Stuss and colleagues (1995; 2005) identify six different types of cognitive processes, each of which are supported by distinct anatomical regions within the PFC: *energising schemata (facilitation and allocation of arousal); monitoring the level of activity within the schema; inhibiting task-irrelevant schemata; adjusting contention scheduling; controlling ('if this then that') logic;* and *task setting* (Stuss et al., 1995; 2005; see also Stuss, 2011).

These predictions form the basis of the ROtman-Baycrest Battery to Investigate Attention (ROBBIA; Stuss et al., 1995). In this battery, Stuss and colleagues (1995) identify a number of specific tasks and measures that recruit and operationalise each of these various processes. Some tasks and measures identify more than one process, but by comparing across tasks, we are able to extract each of the basic elemental processes. These tasks are summarised in Table 4.1.

Table 4.1.

A brief description of the ROBBIA tasks and the hypothesised component processes they recruit (Stuss et al., 1995, 2005).

Hypothesised component processes	ROBBIA Tasks	Description
Energising	Simple Reaction Time (RT)	Detection of, and a response to, a monotonous sequence of stimuli that occur at a slow and relatively infrequent rate over a prolonged period of time
Energising, monitoring, inhibiting	Choice RT	Similar to the Simple RT task, with the additional condition that the stimulus sequence contains both target and non-target stimuli
Energising, monitoring, inhibiting	Prepare RT	Similar to the Choice RT task, with the addition of a preparatory signal presented at variable time lengths preceding the stimulus
Energising, adjustment of contention scheduling, inhibiting	Concentrate	Serial choice reaction time task where responses are made to targets occurring at a rapid rate
Energising, monitoring, task-setting	Count	Counting of stimuli presented at different rates
Energising, monitoring, adjustment of contention scheduling	Divide	Responding to two separate and unrelated tasks that are occurring at the same time
Energising, monitoring	Tap	Simple motor task involving a tap response at a fixed rate both with, and without an external cue
Energising, inhibiting, control of logic	Switch	Switching between two different tasks within the same block of stimuli
Inhibiting, control of logic	No-go	Suppression of a response to a particular stimulus of class of stimuli
Inhibiting, control of logic, task-setting	Suppress	Suppressing a response to a non-target stimulus that shares characteristics with a target stimulus
Energising, monitoring, control of logic, task-setting	Set	Establishing a response mode when response requirement change from one block of stimuli to another

Evaluation. A significant strength of Stuss and colleagues' (1995; 2005) approach was the very simple tasks and the operationalisation of a number of key processes they

propose. This can ultimately aid in the isolation and differentiation of various functions within the prefrontal regions, as well as their precise neural localisation. Accordingly, a "core" processes approach, such as that used by Stuss and colleagues (1995; 2005), may provide a richer source of knowledge compared to other, more generalised neuropsychological tests.

Nonetheless, the model, as it currently stands, is rather broad in its definitions and conceptualisations. For example, each of the three major functions identified in the framework may themselves be further broken down. Indeed, the process of energisation is held to involve a number of components necessary for the *initiation, maintenance,* and *sustained intention* to respond to a stimulus under conditions of low activation or insufficient exogenous arousal (Shallice, Stuss, Alexander, Picton, & Derkzen, 2008; Stuss, 2011; Stuss et al., 1995; 2005). Accordingly, these three processes may make differential contributions to the overall energisation process, as well as the processes of inhibition and attentional monitoring.

Moreover, as discussed in Chapter 1, inhibitory mechanisms also include a number of processes, including the inhibition of prepotent responses, as well as the need to overcome competition and resolve conflict amongst multiple, and entirely appropriate representations (Botvinick et al., 2001). However, the extent to which these processes subserve an overall inhibition mechanism remains unclear. Further, there is debate regarding the way in which inhibitory functions operate – that is, some suggest they operate in a domain-general manner that subserves all material and domains equally, whilst others suggest that inhibitory control is subserved by domain-specific processes (Hamilton & Martin, 2005). This will be discussed further in Chapter 5.

Finally, as is evident from Chapter 1, the concept of *performance monitoring* encompasses a number of different processes. Accordingly, it may be necessary to isolate specific components of monitoring in order to determine how they relate to one another, and with more general processes of cognitive control, such as sustained attention and inhibition.

Introduction to the Current Study

In the current study, our overarching aim was to isolate and identify the specific processes or components that contribute to higher-level cognitive control, and their inter-relationship using tasks that are as "pure" as possible, and which can be performed by a wide range of patients with different types of brain damage. With these considerations in mind, our

starting point was the framework of Stuss and colleagues (1995; 2005). Specifically, we first identified tasks/measures that operationalise each of the three major processes proposed in this model: *energisation, inhibition,* and *performance monitoring.* We then obtained the relevant measures for each participant in our sample, and examined the inter-relationships between our measures.

We selected tasks from Stuss and colleagues' (1995; 2005) ROBBIA test battery that fulfilled the following criteria: i) there was already a body of existing knowledge concerning how various brain-damaged populations perform on the task (where possible including brain tumour populations); and ii) the task contained very few component processes, thus allowing the various proposed components to be teased apart. The following section provides an overview of how each of the three component processes are operationalised in the current study.

Energisation: Activating Schemata and Sustaining Attention

According to Stuss and colleagues (1995), the process of energisation encompasses not just the initiation of an energised response, but also the ability to sustain an intention to respond to a particular type of stimulus (Shallice et al., 2008; Stuss & Alexander 2007; Stuss et al., 1995; 2005). However, for the purpose of the current study, we further divided this set of skills into two subclasses, according to whether they primarily involved the initiation of schemata, hereafter referred to as *activation*; or whether they involve sustaining a response set over time, hereafter referred to as *attentional maintenance.* This subdivision is consistent with recent evidence that suggests distinct anatomical regions may subserve these two sets of processes. For example, superior medial regions have been implicated in the initiation of responses, as evidenced by the following performance patterns: 1) slow response latencies across a number of domains; and 2) significantly slowed response latencies in the first quarter of trials on a simple RT task relative to the second, third, and fourth quarter when compared to healthy control participants (e.g. Alexander et al., 2005; Stuss et al., 2005; see also, Stuss et al., 1998). In contrast, the processes involved in sustaining a response set appear to be mediated by right frontal-parietal regions, including the right anterior cingulate, dorsolateral prefrontal cortex, and parietal cortical regions (Sarter, Givens, & Bruno, 2001).

Activation

Specifically, we operationalise the process of *activation* in terms of latency to respond. One measure of response latency was selected: overall response times on very

simple tasks. The following four tasks from the ROBBIA test battery were used to obtain this latency measure: *Simple RT*, *Choice RT*, *Prepare RT*, and *Concentrate RT*. Each of these tasks were selected due to their simplicity, and also because previous studies have reported slowed response latencies on these tasks in patients with lesions to the superior medial region of the PFC (e.g. Alexander et al., 2005; Stuss et al., 2005). We also included a further measure of latency, which was overall response times from an antisaccade task (described below). Finally, we supplemented these nonverbal measures of response time with a verbal latency measure, specifically naming latencies on the congruent condition of the Stroop task.

Attentional maintenance

The process of *attentional maintenance* (sustained attention) is more difficult to operationalise. In this study, we utilise four different types of measures to operationalise this construct. The first capitalises on the phenomenon known as the *vigilance decrement*. Numerous studies of healthy participants have found that, when participants are presented with a long stream of stimuli during a given trial, response times increase across the duration of a trial (Rueckert & Grafman, 1996; Wilkins, Shallice, & McCarthy, 1987). Indeed, vigilance studies in adults have reported that performance declines considerably over time, particularly for those with difficulty sustaining attention (Rueckert & Grafman, 1996). This phenomenon, known as the *vigilance decrement* (Sarter et al., 2001), has been operationalised in prior studies by comparing mean response times for the first half of simple attention tasks with mean response times for the second half of the trials (see Rueckert & Grafman, 1996). Indeed, in an earlier study, Rueckert and colleagues (1996) utilised this method on individuals with frontal lesions using a series of response time tasks, similar to those in the ROBBIA battery. Patients with right frontal lesions exhibited a disproportionately large increase in latency on the second half of the trials, relative to other frontal groups and controls, suggesting a difficulty sustaining an energised response over time (Rueckert & Grafman, 1996). A more sensitive way to measure the vigilance decrement is to calculate a slope measure that describes the linear latency change across trials, based on the line of best fit. With this consideration in mind, we utilised slope measures for the following tasks from the ROBBIA test battery: *Simple* and *Choice RT*, and the *Concentrate* task. These tasks were selected due to their simplicity and monotonous nature that places high demands on sustained attention processes. We elected not to use the *Prepare RT* task (identical to the *Choice RT* task with the exception of a warning signal, described below) in our slope measure, given the presence of warning tones may have added a confounding effect.

The second measure we used to operationalise attentional maintenance is ISI effects. By introducing long intervals between stimuli, we create conditions that are likely to maximise demands on attentional maintenance – particularly when the stimuli themselves are monotonous (Breckel, Giessing & Thiel, 2011; Davies & Parasuraman, 1982; Warm, Parasuraman, & Matthews, 2008). Consequently, individuals who have difficulty with response set maintenance are likely to show an exaggerated effect of ISI – that is, their response accuracy and/or latency should actually *decrease* as the interval between stimuli increases[15].

The third method we used to operationalise attentional maintenance is by varying warning signals. Stuss and colleagues (2005) created a task wherein target stimuli were preceded by a warning signal, which occurred either 1 second or 3 second prior to stimulus onset ("*Prepare RT* task"). Overall, they found that patients with lesions to the superior medial regions were significantly slower to respond following the 3s relative to 1s warning signal. This performance pattern was attributed to a loss of the intention to respond during the inter-stimulus interval, which is partially mitigated by the warning signal (energisation/ initiation; Stuss et al., 2005). In the current study, patients' response latencies were compared under two conditions: following no warning signal (Choice RT); and following a warning signal presented either one or three seconds prior to stimulus presentation.

The fourth method of operationalising the concept of attentional maintenance was by comparing tasks that placed differing demands on the maintenance of attention, namely, a comparison between the Simple RT task (which only involves one stimulus) and the Choice RT task (which involves four different stimuli and different corresponding responses). Stuss and colleagues (2005) also included this comparison in their study and found that patients with lesions to the superior medial regions of the PFC were significantly slower when the task was particularly monotonous (Simple RT task) compared to when the task contained greater demands of choice (Choice RT task). No other frontal groups showed this effect. In

---

[15] Stuss and colleagues (1995; 2005) offer inconsistent accounts of the "reverse' ISI account in different papers. Following their finding that this pattern of performance was associated with lesions to the right lateral PFC (Stuss et al., 2005), they suggested it may be more indicative of a monitoring impairment than an impairment in energisation/sustained attention. However, a number of fMRI and PET studies have implicated right lateral prefrontal regions in a number of very simple vigilance tasks (Berman and Weinberger, 1990; Cohen et al., 1988; Coull, Frackowiak, and Frith, 1998; Lewin et al., 1996; Nelson et al., 2014; Rueckert & Grafman, 1996). Consequently, it is not surprising that impairments involving these skills may be associated with right lateral frontal damage.

the current study, patients' overall average response latencies on the Simple RT task (SRT) were directly compared with those on the Choice RT task (CRT).

Response inhibition

Operationalising the concept of response inhibition is challenging, because different types of tasks appear to measure different aspects of inhibitory control. Indeed, Nigg (2000) suggests that some tasks primarily assess the ability to focus general cognitive resources towards the current goal or task, and away from a competing goal or task. Others assess the ability to inhibit information within working memory, whilst other tasks assess the ability to inhibit certain kinds of well-learned or reflexive responses, each of which may engage its own specific process (e.g., reflexive saccades, well-learned prepotent response; see Nigg, 2000).

In the current study, we focus on the latter class of inhibitory control tasks: the inhibition of reflexive and/or prepotent responses. Specifically, we examine four types of tasks: 1) The Stroop; 2) the anti-saccade task paradigm; 3) the filter condition of design fluency; 4) and a choice and prepare RT task, in which one stimulus-response pairing occurs more frequently than the others. The Stroop task was selected due to its consistent validation as a test of prepotent response inhibition in neurological tumour populations (e.g. Değerlendirmesi, 2012; Faulkner et al., 2017; Floden & Stuss, 2006; Stuss, Floden, Alexander, Levine, & Katz, 2001; Taphoorn & Klein, 2004), whilst the antisaccade task was selected given it imposes similar demands to the Stroop – namely, suppressing a habitual response to a stimulus in favour of a less well-learned alternative (for a similar argument, see Hamilton & Martin, 2005).

The third task of inhibitory control involves a version of the design fluency paradigm. This task requires participants to draw a series of novel designs by connecting a random array of dots (see Figure 4.2). In the baseline "basic" condition, the array contains black dots, and the participants must utilise these to generate novel designs. In the critical "filter" condition, the array contains both black and empty dots, and the participant must utilise only the *empty* dots in their design, whilst ignoring the *black* dots (see Panel b of Figure 4.2). This second variation arguably introduces an additional element to the task: namely, the ability to ignore a class of stimuli that was previously important to the task, in favour of a new class of stimuli. Specifically, we compare the proportion of designs generated during the filter condition of

the design fluency task, to that of the basic condition, which does not make the same demands.

Our fourth set of tasks were based on the Choice and Prepare RT tasks from the ROBBIA test battery; in these tasks, the participant must decide whether each stimulus is a target or a non-target. The target stimulus (e.g., the letter A) appears 25% of the time, whereas the non-target stimuli (e.g., the letters, B, C and D) occur in the remaining 75% of trials. In other words, the "dominant" response is "non-target", and this dominant response must be suppressed in those rare instances in which a target stimulus (the letter A) appears. In this way, the task parameters create one "prepotent" response which must be overcome in order to make the less frequently occurring responses (for a similar argument, see Stuss et al., 2005)

Performance Monitoring

In our framework, we conceptualised performance monitoring as a set of processes that enable the person to adjust their behaviour in response to internal or external feedback. External feedback can include sudden changes in the environment, or feedback about how successful our current strategies are in relation to the current goal. Internal feedback can include evaluating a planned behaviour or response to ensure it is in line with current goals before implementing it. To operationalise this concept, we examined two phenomena: 1) post-error slowing, and 2) perseverative and inappropriate errors on open-ended tasks.

Our first measure, post-error slowing (PES), refers to the tendency of healthy participants to exhibit slowed response times immediately following the commission of an error (Botvinick et al., 2001; Notebaert, Houtman, Van Opstal, Gevers, Fias & Verguts, 2009). This phenomenon suggests that the participant has monitored their performance and consequently adjusted their behaviour in response to the previous error - that is, response latencies following an error-trial can be used as a marker for effective monitoring, as we define it here, and in the literature (e.g., Botvinick et al., 2001; see also Bogte, Flamma, van der Meere, & van Engeland, 2007; Kerns et al., 2005; Notebart et al., 2009).

To measure the construct of post-error slowing, we will compare response latencies immediately following the commission of errors to response latencies following post-correct trials on six tasks (two choice tasks; an antisaccade task; a concentrate task; and two BLAST subtests: Stroop and Picture-Word Verification). These tasks were selected for the following two reasons: first, they were likely to produce more errors than the remaining BLAST

subtests, and second, errors were likely to be more "noticeable" and overt. We reasoned that individuals with an impaired monitoring system will fail to regulate their behaviour following the commission of an error, as evidenced by minimal differences in response latency between post-error and post-correct trials.

Our second measure of performance monitoring focuses on specific classes of errors in more open-ended tasks, including perseverative errors and responses that fail to comply with task instructions. As discussed above, it has been suggested that perseverative errors reflect a failure to correctly and continuously monitor internal representations of response plans, in order to ensure they are in line with task requirements (Reverberi et al., 2005; Shao et al., 2014). Specifically, many spontaneous generation tasks, such as verbal fluency include the requirement that a person should not generate the same exemplar more than once. Therefore, in order to fully comply with the task instructions, a person must constantly update and monitor their planned responses for potential duplicates. Here, we will examine rates of perseverations on the following tasks of unconstrained and spontaneous generation: verbal fluency and design fluency. A high proportion of perseverations will be interpreted as a marker of impaired monitoring.

Specific Task Measures and Predictions

Table 4.2. outlines each of the four posited cognitive skills and the key performance measures that contribute to their operationalisation.

Based on the theoretical framework proposed by Stuss and colleagues (1995; 2005; see above), we make the following predictions regarding the inter-relationships between our component measures. First, if our assumption is correct, that cognitive control can be decomposed into the four core skills we identified above, then we would predict that individuals' measures for each these skills will not correlate significantly with one another. Further, we would expect to see at least some cases where two or more skills dissociate (that is, individuals score highly on one skill but not the other).

Table 4.2 also summarises our predictions as to the groups most likely to be impaired on each core skill, based on the primary localisation of their tumour (see also Chapter 1 for discussion).

Table 4.2.

Each of the four posited cognitive skills, the key performance indicators, and the corresponding predictions of each skill

Cognitive Skill/ Key performance indicators	Key performance indicator	Prediction
Activation/ Energisation		
Slowed response latencies on a range of simple verbal and nonverbal tasks	Slowed response latencies on four RT tasks from the ROBBIA (Simple RT, Choice RT, Prepare RT, and Concentrate task). Slowed response latencies on the antisaccade task. Slowed naming latencies in the congruent condition of the Stroop task	Based on the hypothesis that this core skill is supported by bilateral regions in the superior medical PFC, we predict that low scores will be associated with primary lesions within the left or right PFC. We do not predict a difference between left and right PFC groups.
Attentional Maintenance		
Disproportionately slow performance when inter-stimulus intervals (ISIs) and warning intervals are long;	On simple, button press tasks, slower response latencies as the trials progress On button press tasks with variable ISIs, slower responses to trials following longer,	Based on the hypothesis that attentional maintenance is supported by right PFC, we predict that low scores on the attentional maintenance measure will be associated with primary lesions to the right PFC.

this effect is attenuated as the task progresses (vigilance decrement). Slower latency performance on more monotonous, compared to complex, trials

when compared to those following shorter ISIs.

On an RT task that incorporates pre-trial warnings (Prepare RT) slower responses on trials following a 3s warning than on trials following a 1s warning.

Paradoxically longer response latencies on Simple than on the Choice RT tasks.

Inhibition

Disproportionately slow performance on prepotent items that cause a lot of conflict and require the individual to overcome their dominant response

On the Stroop task, abnormally slow naming latencies on incongruent items (where word name and colour name do not match), relative to congruent items.

Abnormally high error rates on the antisaccade task.

In choice reaction time tasks (the Choice RT task and the Prepare RT task), abnormally prolonged response latencies for stimulus-response pairings with a low rate of occurrence.

Based on the hypothesis that verbal inhibition is supported by left PFC, we predict that low scores on this measure will be associated with primary lesions to the left PFC. However, it remains possible that lateralisation may also depend upon the nature of the response required – for example, verbal vs. nonverbal; (we consider this issue in more detail in the next chapter) – therefore, based on previous observations of right frontal involvement in non-verbal tasks (e.g. Aron et al., 2003; Floden & Stuss, 2006; Picton et al., 2006), we also predict that low scores on these non-verbal measures will be associated with primary lesions to the right PFC

On the design fluency task, a marked
increase in errors on the "Filter" condition
relative to the "Simple" condition.

Performance Monitoring

The commission of multiple errors and difficulty monitoring self-performance	Abnormally high rates of preservative and/or inappropriate errors on the Design Fluency task Abnormally high rates of preservative and/or inappropriate errors on Verbal Fluency tasks Absence of post-error slowing on button press RT tasks (no differentiation between post-error and post-correct trials)	Based on the hypothesis that performance monitoring is supported by right PFC, we predict that low scores on this measure will be associated with primary lesions to the right PFC.

Method

Participants

The sample of tumour patients was identical to that described in Chapter 2. All 28 neurological tumour patients completed these tasks as part of a single study, over the course of a single testing session (see Chapter 2 for details of recruitment and consent procedures).

An additional sample of 23 healthy controls was tested on five tasks, for which no published control data was available. The tasks were: *Simple Reaction Time*; *Choice Reaction Time; Prepare Reaction Time; Anti-saccade;* and the *Concentrate Task.*

Control participants were grouped according to age: 13 were aged between 26 and 50 (mean age: 36 years, range = 26.3 – 49.5 years) and consisted of six males and seven females; and 10 were aged 51 years and over (mean age: 64 years, range = 51.3 – 78.2 years), and consisted of four males and six females. Using SPSS Software, an independent samples t test revealed that our two groups did not differ significantly in age.

In order to exclude participants with any significant neurological disease/injuries, visual or hearing deficits, or motor control impairments, participants completed a brief self-report Neurological Status questionnaire prior to the study proper (see Appendix E). The consent form also directly asked participants about any current problems in the above categories (see Appendix E). No participants were excluded on this basis. Control participants were recruited through the following methods: i) through the Victoria University IPRP Programme (this programme allows first year psychology students to participate in experiments in exchange for course credit); ii) through a pre-existing register of older healthy controls that had indicated their willingness to be contacted about future studies in the neuropsychology lab at Victoria University of Wellington; and iii) through community advertising. The recruitment material specified that participants needed to speak English as their native language, and participants were asked to confirm this verbally at the start of their first session. Prior to taking part in the study, all participants were required to read through an information sheet and sign a consent form (see Appendix E).

Testing of control participants occurred either at Victoria University of Wellington or at the participants' own homes. Participants recruited through IPRP received course credit, whilst all others received a monetary contribution in exchange for their participation ($10 if the examiner travelled to their house, or $20 if the participant travelled to Victoria University

of Wellington). The testing of healthy control participants was approved by Victoria University of Wellington's School of Psychology Ethics Committee.

Materials and Procedure

Tasks. Eight tasks were administered to participants. The Simple Reaction Time (SRT) task, Choice Reaction Time (CRT) task, Prepare Reaction Time (PRT) task, and the Concentrate task were all adapted from Stuss and colleagues (2005), with each forming part of the ROtman-Baycrest Battery to Investigate Attention (ROBBIA; Stuss, et al., 1995). Several other tasks used here were adaptations of previous tasks used in the literature; details are given below. The verbal fluency task and the Stroop task were both administered as part of the BLAST protocol, the procedure for which is described in Chapter 2.

The following six tasks were administered using PsyScope software (Cohen et al., 1993) on a MacBook Pro laptop that also collected response latencies: Simple RT, Choice RT, Prepare RT, Concentrate, Antisaccade, and the BLAST Stroop task. The BLAST verbal fluency task was administered orally (see Chapter 2), whilst the design fluency task was administered in a pen and paper format according to the standard procedures outlined in Delis, Kaplan and Kramer (2001). The materials and procedure for each of the remaining six tasks are summarised below.

Simple Reaction Time task. In this task, each trial involved presenting the capital letter 'A' in black ink on the center of a white screen, in Times New Roman, size 72 font. There were a total of 40 trials. Inter-stimulus intervals (ISI), defined as the period from the last response until the onset of the next stimulus, were 3s, 4s, 5s, 6s, & 7s[16]. Each ISI occurred eight times in a pseudorandom, blocked order a total of five times, with the limit that no two successive trials had the same ISI. This order remained constant for all participants.

During this task, participants were required to press the space-bar on the keyboard as soon as the 'A' stimulus appeared. This response removed the stimulus from the screen, and initiated the next trial. Participants were told to respond as quickly, and as accurately as possible. Two practice trials, with ISI's of 3s and 7s respectively, preceded the 40 experimental trials. Feedback was provided throughout the practice phase, and the examiner

---

[16] Technically, these are response-stimulus intervals, since the fixed duration is timed relative to the onset of the participant's *response* to the preceding trial, not from stimulus onset to stimulus onset. However, we have continued to use the term to maintain consistency with Stuss and colleagues' (2005) terminology.

ensured participants were in an adequate viewing position before the experimental trials commenced. Participants were informed that they were free to stop at any time throughout the duration of the task.

The Choice Reaction task. This task was identical to the Simple Reaction Time task with the exception that there were four stimulus letters: the letters 'A', 'B', 'C', and 'D' each appeared 10 times in a pseudo randomised blocked order across 40 trials. Participants were required to press 'X' on the keyboard when they saw the stimulus '*A*', and 'M' on the keyboard each time they saw any of the other stimuli ('*B*', '*C*' or '*D*'). The X and M keys were labelled '*A*' and '*Other*' respectively to minimise any confounds of memory. Four practice trials preceded the 40 experimental trials, during which all four stimulus alternatives were presented and feedback was provided.

The Prepare task. This task was identical to the Choice Reaction Time task, with the following three exceptions. First, a warning stimulus preceded each trial; a single tone lasting 200ms was presented either 1s or 3s before the onset of the letter stimuli. The 1s and 3s warning tones each occurred 20 times in a pseudo-randomised, blocked order of 40 trials, with each warning tone occurring five times for the '*A*' stimuli, and five times for the '*B*', '*C*' and '*D*' stimuli. Second, the ISIs were set at 4s, 5s, 6s and 7s. Third, participants were told to expect a warning signal prior to each letter stimulus, which may facilitate their responding.

The Concentrate task. This task was identical to the Choice Reaction time task, with two exceptions. First, the stimulus letters were replaced with four colour words: *red, yellow, blue,* and *green*. Second, the inter-stimulus interval was kept constant at 500ms. Each target word was presented in its respective ink colour, *red, yellow, blue,* and *green,* in Ariel, size 72 font in the centre of a white screen. The four target words appeared 10 times in fixed blocked order over 50 trials; each block contained the four target words. This order remained constant for all participants. Participants were required to press key 1 each time the word "*red*" was presented, key 2 each time the word *"blue"* was presented, key 7 each time the word *"yellow"* was presented, and key 9 each time the word "*green*" was presented. Each target key, as well as the keys directly above, were labelled in their respective colours to minimise any confounds of memory. Feedback was provided over eight practice trials which preceded the 40 experimental trials; during this period, each of the four colour alternatives were presented two times.

The Anti-saccade task. This task was adapted from Hamilton and Martin (2005; see also Roberts, Hager & Heron, 1994). During the task, a fixation cross was presented in the centre of the screen for 1500ms (unlike Hamilton & Martin, 2005, we did not manipulate the duration of fixation time). Following fixation, a cue appeared 3.4 inches from the fixation cross in the form of a small black square, which appeared on either the left or right-hand side of the screen. After 175ms, the cue disappeared, and a black arrow was presented on the opposite side of the screen from the cue. Again, this arrow appeared 3.4 inches from the fixation cross and pointed to one of four directions: the far left side of the screen, the far right side of the screen, on the far left side pointing up, and on the far right side pointing up. After 150 ms, the arrow was replaced by cross-hatching which remained visible in the centre of the screen until the participant responded. Figure 4.1. illustrates an example trial in the antisaccade task. Each of the four arrow orientations appeared ten times in a pseudo randomised, blocked order over 40 trials.
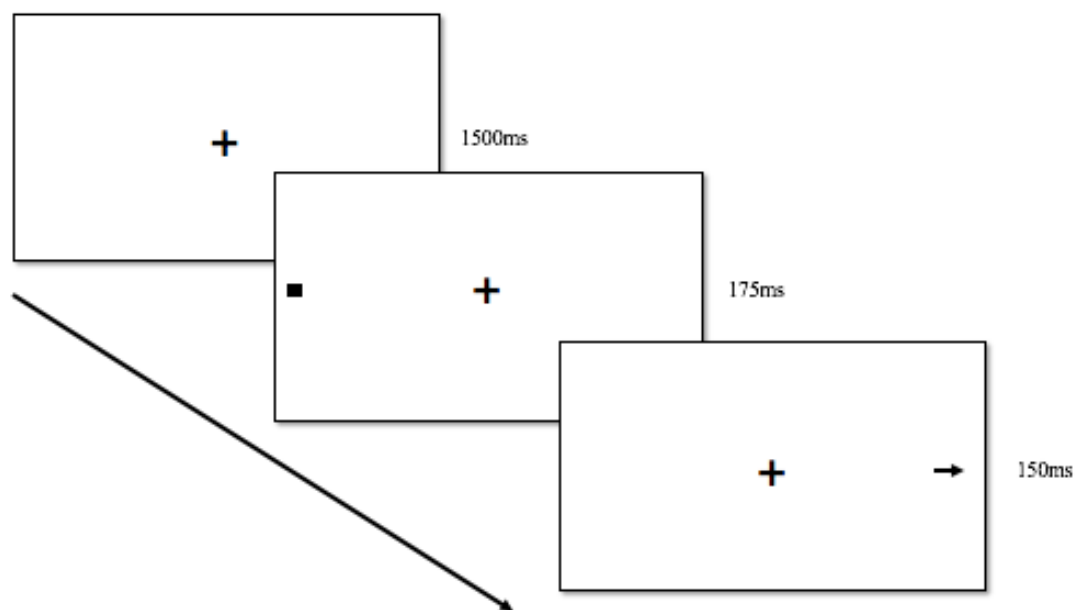
Figure 4.1. An example of a typical trial from the antisaccade task. In this example, the cue is presented on the left-hand side of the screen, followed by the target arrow, which is pointed to the right-hand side of the screen. The participant must ignore the cue (and continue looking at the fixation cross) and then press the arrow that corresponds to the direction of the target arrow, which in this example would be the right arrow.

In this task, participants were asked to ignore the cue, and only attend to the direction of the arrow. When the arrow was pointing up, individuals were required to press the "up"

arrow key on the keyboard, regardless of which side of the screen the arrow appeared on. When the arrow was pointing to the left or right, they were required to press the left or right arrow keys respectively. The response initiated the start of a new trial.

Eight practice trials preceded the 40 experimental trials, in which the examiner provided detailed feedback, and highlighted and corrected any errors. During the practice phase, each of the four arrow directions appeared two times.

Response latencies for correct responses were measured from the onset of the target arrow, until the onset of a response. Incorrect responses were not recorded in the overall latency data. Finally, data was not further analysed for any participant who met the following criteria: a) their responses exhibited a systematic bias towards one side of space (either left or right; such a pattern could indicate a primary visual perception deficit or an attentional bias to one side); and b) they produced errors on more than 97% of trials (again, such a pattern may indicate a visual difficulty). Response latencies were analysed in the same way previously described. First, prior to data analysis, latency data were Winsorised and trimmed of any outliers: the longest response latency was replaced by the second longest response latency, and the second longest response latency was replaced by the third longest response latency. Then, any latencies that lay two and a half standard deviations from the mean were removed.

The Design Fluency task. The Design Fluency task was taken from the Delis-Kaplan Executive Function Scale (D-KEFS; Delis, Kaplan, & Kramer, 2001). Standard administration and scoring protocols were followed. The task has three conditions: *basic, filter* and *switch;* each was administered in this order during the course of a single session. In each of the three conditions, participants were shown a sheet of paper containing an array of dots. They were required to produce novel designs by connecting the dots using four straight lines. Their instructions were that each line should touch both its origin and destination dot, the lines can cross each other, and each design had to be unique. Examples of some possible designs are shown in Figure 4.2. Participants did not have to draw the line using one continuous stroke; they could lift the pen from the paper at any time. These instructions were presented both orally, and in written form.

Figure 4.2a *Basic condition*

| Blank design | Acceptable design | Perseverative error | Inappropriate error |
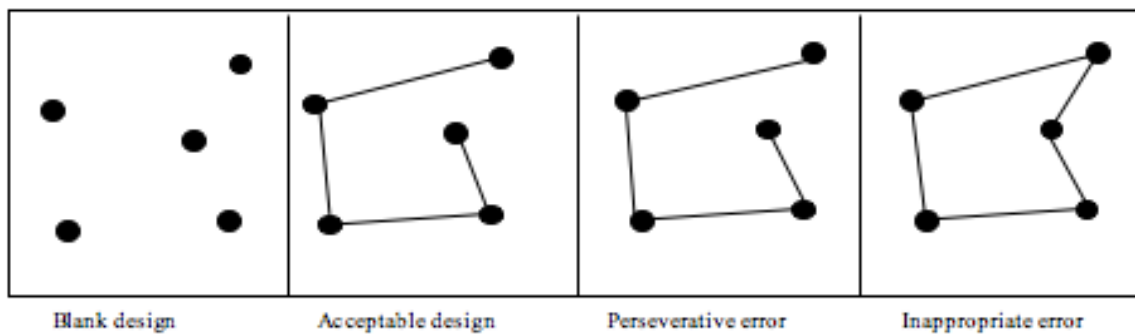
Figure 4.2b *Filter condition*

| Blank design | Acceptable design | Perseverative error | Inappropriate error |
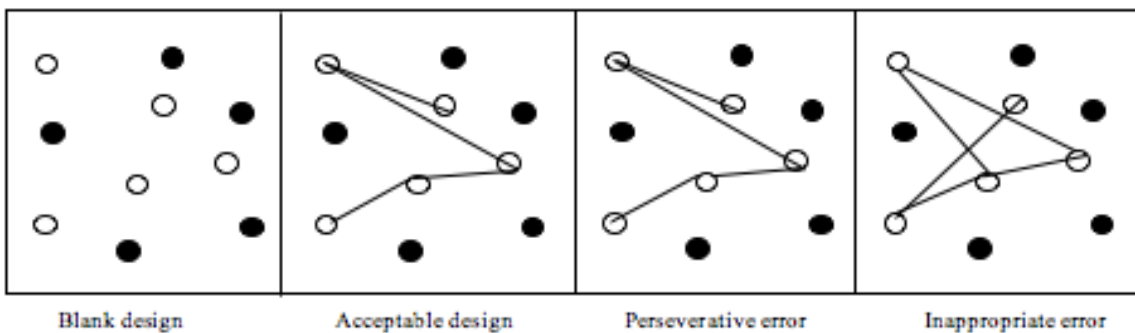
Figure 4.2c *Switch condition*

| Blank design | Acceptable design | Perseverative error | Inappropriate error |
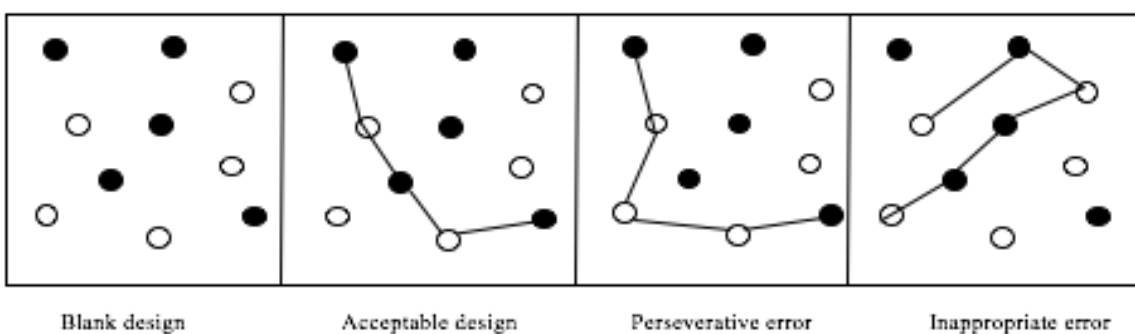
Figure 4.2. Examples of each of the conditions in the design fluency task. Participants were required to draw novel designs using four lines. Examples of inhibitory control deficits are presented.

There were three conditions. In the *basic* condition, each square contained an array of five black dots. Patients were required to produce novel designs by connecting dots using only four straight lines (Figure 4.2a). They were allowed 60s to complete the task. Once the time limit was reached, the examiner administered the next condition. In the *filter* condition, each square contained an array of black and empty dots. This time, patients were required to connect only the empty dots, again using four straight lines (Figure 4.2b), thus 'filtering' out

172

the dots they had previously connected in the basic condition. Again, a time limit of 60s was imposed. Following this, the *switch* condition was then administered. Here, each square contained a new array of black and empty dots. Patients were required to make novel designs by switching between black and empty dots (figure 4.2c). Again, only four lines could be used. Individuals could start with either a black or an empty dot. The same general rules applied throughout the three conditions. Again, patients were given 60s to complete the task.

At the commencement of each condition, patients were shown a practice page containing three example arrays. The examiner firstly demonstrated a suitable design, followed by individual practicing a design of their own. Detailed feedback was then given. Following each practice trial, participants were presented with the experimental condition, containing 35 arrays per condition. The rules were then repeated verbally. Individuals were instructed to draw as many novel designs as possible within 60 seconds. A stopwatch was used to time each condition. Patients were permitted to finish any design in progress when the time limit was reached. During the experimental phase, the examiner was able to prompt the participant if they drew three consecutive incorrect or perseverative designs (e.g. *"remember to use only four lines"*).

Error types were classified as either perseverative or inappropriate errors. Perseverative errors were defined as any repetitions of the same design within the same condition (see Figure 4.2). Inappropriate errors were defined as those which were inconsistent with the task instructions; these included designs that contained fewer than, or more than four lines (see Figure 4.2).

Calculation of the key cognitive measures

As summarised in Table 4.3. below, we used this battery of tasks to obtain measures of our four proposed cognitive skills: *activation, attentional maintenance (sustained attention), inhibitory control* and *performance monitoring*. Below are further details of how the various measures were calculated. As seen in Table 4.3., each of our selected individual measures were combined in such a way that we had four aggregated measures for each of our four posited control skills.

Table 4.3

Each of our four hypothesised core skills, and the measures used to operationalise them. See text below for more detailed information about how each measure was calculated.

Cognitive Skill	Key Measure	Final Skill Score submitted to analysis	Formula
Activation	a) *Reaction time speed*: Average latencies in four different button press tasks (Simple RT, Choice RT, Prepare RT and Concentrate RT), expressed as a z score relative to the entire patient sample (see below for details).	Measure a), expressed as a z score relative to the entire patient sample (that is, using the mean and standard deviation of the patient sample).	Mean (a, b, c)
	b) Average latencies on congruent trials in the Stroop task, expressed as a z score	Measure b), expressed as a z score relative to the entire patient sample.	
	c) Average latencies in the antisaccade task, expressed as a z score	Measure c), expressed as a z score relative to the entire patient sample.	
Attentional Maintenance	d) For simple RT, choice RT, prepare and Concentrate tasks, the change in latencies over the course of the trial (as measured using a slope paradigm) expressed as a z score	Measure d), expressed as a z score relative to the entire patient sample.	Mean (d, e, f, g)

	e) For simple RT and choice RT, the difference in average response latencies for long ISI's (6 & 7s) and short ISI's (3 & 4s)	Measure e), expressed as a z score relative to the entire patient sample.	
	f) The difference in average latencies on the choice reaction time task and the simple reaction time task, expressed as a z score	Measure f), expressed as a z score relative to the entire patient sample.	
	g) On the prepare task, the difference in response latencies on trials following a 1s warning tone and those following a 3s warning tone, expressed as a z score.	Measure g), expressed as a z score relative to the entire patient sample.	
Inhibitory Control	h) On the choice RT and prepare RT tasks, the average response times for letter A trial, expressed as a proportion of the average response time for the remaining trials (B,C, or D), expressed as a z score	Measure h), expressed as a z score relative to the entire patient sample.	Mean (h, i, j, k)
	i) Average accuracy on the antisaccade task, expressed as a z score	Measure i), expressed as a z score relative to the entire patient sample.	
	j) On the Stroop task, the average naming latencies for incongruent trials as a proportion of the average	Measure j), expressed as a z score relative to the entire patient sample.	

naming latencies for congruent
trials, expressed as a z score

	k) The total number of correct designs produced on the Filter condition as a proportion of the total number of correct designs produced on the Basic condition, expressed as a z score	Measure k), expressed as a z score relative to the entire patient sample.	
Performance Monitoring	l) On the Choice RT, prepare RT concentrate, antisaccade, Stroop and picture-word verification tasks, the difference in average response latencies on post-error and post-correct trials, expressed as a z score	Measure l), expressed as a z score relative to the entire patient sample.	Mean (l,m,n)
	m) The total number of errors on the design fluency task as a proportion of the total number of correct designs, expressed as a z score	Measure m), expressed as a z score relative to the entire patient sample.	
	n) The total number of errors on the verbal fluency task as a proportion of the total number of correct responses, expressed as a z score	Measure n), expressed as a z score relative to the entire patient sample.	

Activation (*processing speed*): Our average latency measures were calculated as follows. The first measure (Measure a in Table 4.3) was obtained by calculating the global average response time for each participant on each of the following tasks: i) Simple reaction time; ii) Choice reaction time; iii) Prepare reaction time; and iv) Concentrate. Each individual's response latencies for each task were first trimmed of outliers according to the procedures set out above and in Chapter 2, and then averaged. Then, a grand average was calculated for each patient from the means for all four tasks.

Measure b was calculated from the (trimmed) latency data from Stroop task, which was administered as part of the BLAST assessment (Chapter 2). Only response latencies to congruent trials were used in this calculation.

Finally, Measure c was calculated from response latencies on the antisaccade task; latencies were first trimmed of outliers in the same manner as previously outlined, and then a grand mean was calculated for each individual participant.

Attentional maintenance (*or Sustained attention*). Measure d in Table 4.3 above was calculated using data from the following four tasks: Simple RT, Choice RT, Prepare RT, and Concentrate. We first calculated a slope measure for each task for each person that expressed the rate of change in latencies across successive trials (using trimmed latency data). We then combined these four slope estimates to create a global average slope measure for the four tasks for each individual participant.

Measure e was obtained using latency data (trimmed of outliers) from the Simple and Choice RT tasks. We first calculated response latencies for each participant on long ISI's (6s & 7s) and short ISI's (3s & 4s) for each task. For each task, we then created a difference score between long and short ISI's. Finally, we then calculated an average difference score for each participant, and averaged these two scores to create an overall difference score.

Measure f was obtained using latency data (trimmed of outliers) from the Choice RT and Prepare RT tasks. We first calculated the global average response time for each participant on each of the two tasks. We then calculated the difference between these two measures.

Measure g was calculated using data from the Prepare task: we calculated average latencies for responses following the 1s and 3s warnings respectively (using trimmed data), and then obtained a difference score that expressed the difference in response times between the 1s and 3s conditions.

Inhibitory Control. Measure h in Table 4.3 above was obtained using data from the Choice RT and Prepare RT tasks. We first calculated the average response latencies (trimmed of outliers) for each task for each person in response to the "A" stimulus (which occurred in 25% of trials), and the corresponding figure for the remaining stimuli combined. We then calculated a score that expressed the former value as a proportion of the latter value. We did this separately for each task. Then finally, we averaged these two scores to create a grand proportion score.

Measure i was obtained using data from the antisaccade task: the number of correct trials was first calculated (out of 40), and then a grand mean was calculated for each individual participant.

Measure j was obtained from the BLAST Stroop task. We first calculated the average response latencies (trimmed of outliers) for each person in response to the "congruent" stimuli and the "incongruent" stimuli. We then calculated a score that expressed the latter value (average latencies on incongruent trials) as a proportion of the former value (average latencies on congruent trials).

Measure k was obtained from the Design Fluency task. First, we calculated the total number of designs produced in 60s on the "basic" condition and the total number produced on the "filter" condition. We then calculated a difference score between these two conditions.

Performance Monitoring. Measure l from Table 4.3 was obtained from the following six tasks: the Choice RT task, the Prepare RT task, the Concentrate task, the antisaccade task and two tasks from the BLAST: Stroop and Picture-Word Verification. We first calculated the average latencies (trimmed of outliers) for each task for each person for trials that followed an error, and those for trials that followed a correct response. We then calculated a score that expressed the difference in latencies on these two types of conditions (error trials vs non-error trials). We did this separately for each task. Finally, we averaged these two scores to create a grand mean difference score, which we referred to as "post-error slowing".

Measure m was based on the Design Fluency task. We calculated the frequency of perseverations and inappropriate errors (combined score) as a proportion of the total number of designs produced on each condition in the design fluency task and then combined these scores. Perseverative errors were defined as exact repetitions of responses (e.g. producing the

same design twice within a condition). Inappropriate errors were defined as any response that deviated from the task instructions (for example, three or five lines instead of four).

Measure n was based on the verbal fluency task. We calculated the frequency of i) perseverative errors, defined as exact repetitions of responses or morphological variants (e.g. *"eat"* followed by *"eating"*), and ii) inappropriate errors, defined as any response that deviated from the task instructions (for example, proper nouns). These two frequencies were calculated as a proportion of the total number of words produced on each condition and then combined into an overall score.

Each of the measures (a-n) were then expressed as z scores, using the formula detailed in Chapter 2. This method expressed each patients' scores relative to the patient cohort as a whole. Data for controls was treated in exactly the same manner – that is, each raw score was subtracted from the mean for the patient group for that measure and divided by the standard deviation for that group on that measure. As noted in Chapter 2, the primary purpose of standardising scores in this study was not to provide an estimate of where each patient "stood" relative to heathy participants, but rather to ensure all measures used comparable units and had a similar spread. For these purposes, it is better to use the patients themselves as the reference sample, rather than the relevant control sample, because the latter often have a very limited range (for example, some measures were subject to ceiling effects). Following this standardisation, scores were then combined into the appropriate formulae (see Table 4.3) to derive a total score for that skill. This standardisation process ensured that each of the measures were weighted equally, regardless of the measurement scale.

Results

Table 4.4 presents each patient's results for each of the four aggregate skill scores: *activation, attentional maintenance, inhibition,* and *performance monitoring*. Cells highlighted in blue indicate that the participant's score was significantly impaired relative to age-appropriate controls, according to Crawford and Howell's (1998) modified t test. As discussed in Chapter 2, this form of comparison was selected as it treats the control sample as statistics rather than parameters, and has been confirmed as a robust method that controls for skewedness in the data, as well as controlling for type 1 error rate, regardless of the size of the control sample (see Crawford & Garthwaite, 2006). On our individual measures of performance monitoring, it was not possible to examine significant impairment relative to healthy controls, given the control samples produced insufficient errors to obtain a reliable average and standard deviation. These cells are highlighted in grey.

As seen in Table 4.4., 43% of patients exhibited significant impairment on at least one aggregated core skill measure relative to their appropriate control group. This relatively high rate of impairment suggests that a core skills approach may indeed be sensitive at detecting higher-level control processes in neurological tumour populations. Overall, *activation* appeared to be the most sensitive core skill (21% of patients impaired), followed by *attentional maintenance* (14% of patients impaired). Conversely, *inhibitory control* was less sensitive (7% of patients impaired), which may be attributed to the combination of both verbal and non-verbal measures that may have washed out any significant effects. Importantly though, despite the sensitive of the core skills approach, the rates of significant impairment appeared to be non-selective amongst the tumour localisations.

As predicted, there was evidence of dissociations between our skill scores. For example, patient's TD and LC were significantly impaired on *attentional maintenance*, yet showed intact performance on the remaining skill measures. Conversely, patient GP was significantly impaired on the *activation* measure, yet did not show evidence of any other impairments. This pattern of performance is broadly consistent with a non-unitary model of cognitive control.

Table 4.4

Each patient's z scores on each of our aggregated measures of cognitive control. Light blue indicates scores that are significantly impaired according to Crawford and Howell's (1998) modified t test. Z scores for performance monitoring were unable to be statistically compared to the relevant control sample, and therefore these cells are highlighted in grey.

	Attentional Maintenance Average	Activation Average	Inhibition Average	Monitoring Average
Left Frontal				
BP	0.067	1.032	-0.129	0.095
RF	-0.414	-0.618	0.539	0.691
TD	-0.464	0.749	-0.191	0.478
LC	-0.453	0.430	0.036	-0.003
DP	0.341	-0.019	-0.231	0.366
Average	**-0.184**	**0.315**	**0.005**	**0.325**
SD	**0.368**	**0.651**	**0.316**	**0.283**
Left Temporal				
CA	-0.293	-0.842	-0.318	-0.622
SH	-0.582	-0.211	-0.179	-0.497
GM	-0.335	-0.137	0.118	-0.408
KW	-0.152	1.025	0.086	-0.909
Average	**-0.341**	**-0.041**	**-0.073**	**-0.609**
SD	**0.179**	**0.778**	**0.211**	**0.218**
Right Frontal				
NH	0.565	0.048	-0.005	0.635
DF	0.092	0.241	-0.030	0.984
BS	-0.022	-0.258	0.175	-0.062
LW	-0.175	-1.097	-0.419	-0.038

GP	2.152	-1.191	0.089	-1.731
RS	0.038	0.283	-0.158	-0.619
SN	0.246	0.201	-0.813	0.981
WR	0.038	0.388	0.143	-0.735
Average	**0.367**	**-0.173**	**-0.127**	**-0.073**
SD	**0.754**	**0.630**	**0.336**	**0.942**
Left Parietal				
AVG	-0.469	0.423	0.055	0.250
BD	-0.424	-0.159	0.090	0.190
CR	-0.070	-0.254	-0.222	-0.005
NOH	0.240	-1.061	-0.149	0.259
KG	0.076	0.636	-0.086	0.888
SO	-0.279	-2.755	0.670	-0.185
Average	**-0.154**	**-0.528**	**-0.060**	**0.233**
SD	**0.284**	**1.242**	**0.322**	**0.364**
Right Posterior				
SG	-0.257	0.071	-1.221	0.593
AEK	-0.281	0.067	-0.149	0.497
JB	-0.231	0.591	0.049	-0.696
PC	0.225	0.503	0.510	-0.141
Average	**-0.136**	**0.308**	**-0.203**	**0.064**
SD	**0.241**	**0.278**	**0.733**	**0.602**
AE	0.823	0.541	1.292	-0.059

Each patient's scores for each of the individual component measures are presented in Appendix F (Table F.1). All but one of the patients in our sample (96%) showed a significant impairment on our individual component measures. The only patient who was not significantly impaired was the youngest patient, BP, who presented with a left frontal meningioma. The high rate of significant impairment on our individual component measures again highlights the sensitivity of a core skills approach in neurological tumour populations. As seen in Appendix D, our individual measures of *activation* appeared to be the most sensitive at detecting significant impairment, however again, the poor scores did not appear to be concentrated in any specific localisation group for any particular measure (see Appendix F, Table F.1).

Turning now to the first objective of this Chapter, which was to examine the inter-relationships between each of our aggregated skill measures. To do this, we performed a Pearson's correlation using SPSS Software. Table 4.5 presents the uncorrected Pearson's r values for the four major cognitive skill scores.

Table 4.5

Pearson's correlation coefficients for each of our four aggregated cognitive control skills.

		Attentional Maintenance	Activation	Inhibitory Control	Performance Monitoring
Attentional Maintenance	Correlation	-	-0.118	0.185	-0.319
	P Value		*.549*	*.346*	*.098*
Activation	Correlation		-	-0.090	0.171
	P Value			*.648*	*.383*
Inhibitory Control	Correlation			-	-0.238
	P Value				*.223*

*$p<.05$

As can be seen from the Table, there were no significant correlations between any of our four putative cognitive control skills. Such findings are consistent with our original hypothesis that these various aspects of cognitive control can be fractionated into functionally distinct processes. However, of course, this hypothesis would need to be tested in a larger sample before we could unequivocally confirm it.

Following this, based on the assertion that inhibitory control may operate in a modality-specific manner localised to left frontal and right frontal regions (see Chapters 1 and 5), we decomposed our inhibitory control measure into verbal and non-verbal domains. Table 4.6. presents the summary data for each patient on our verbal and non-verbal measures of inhibitory control.

As is evident from Table, elucidating inhibitory control to a modality-specific level reveals more significant impairments. Specifically, 18% of patients performed significantly more poorly than controls on the non-verbal measure of inhibitory control. Conversely, no patients were significantly impaired on the verbal inhibition measure (the difference in response latencies on the Stroop task); this likely reflects the poor performance of control participants on this task.

Table 4.6.

Individual z scores for each of the aggregated measures of verbal inhibitory control and non-verbal inhibitory control. Blue cells represent significant impairment (p< 0.05) relative to the appropriate control sample using Crawford and Howell's (1998) modified t test.

Patient	Total Verbal Inhibition	Total Non-Verbal Inhibition
Left Frontal		
BP	-1.136	0.207
RF	0.219	0.645
TD	-0.978	0.071
LC	-0.867	0.337
DP	-0.799	0.053
Average	*-0.712*	*0.263*
SD	*0.536*	*0.243*
Left Temporal		
CA	-0.291	-0.328
SH	0.921	-0.546
GM	-0.573	0.349
KW	-0.467	0.270
Average	*-0.102*	*-0.064*
SD	*0.692*	*0.441*
Right Frontal		
NH	0.771	-0.264

DF		-0.030
BS	-0.160	0.286
LW	-0.479	-0.399
GP	0.916	-0.187
RS	-0.189	-0.148
SN	0.714	-1.322
WR	0.311	0.059
Average	*0.269*	*-0.251*
SD	*0.551*	*0.481*
Left Parietal		
AVG	-1.166	-0.463
BD	2.091	0.911
CR	0.377	0.422
NOH	-0.295	0.100
KG	-0.193	0.050
SO	2.375	0.182
Average	*0.531*	*0.200*
SD	*1.410*	*0.453*
Right Posterior		
SG	-0.764	-1.374
AEK	-0.035	-0.187
JB	-1.607	0.601

PC	-0.542	0.860
Average	*-0.737*	*-0.025*
SD	*0.655*	*1.003*
Undifferentiated		
AE	1.845	1.108

To assess the relationship between verbal and non-verbal indices of inhibitory control and the remaining skills scores, we again performed a Pearson's correlation using SPSS Software. Table 4.7. displays the uncorrected Pearson's r values for each measure.

Table 4.7.

Pearson's correlation coefficients for our measures of verbal and non-verbal inhibitory control and the remaining cognitive control skills.

		Attentional Maintenance	Activation	Monitoring	Non-verbal Inhibition	Verbal Inhibition
Attentional Maintenance	Correlation	-	0.118	-0.319	0.074	0.264
	P Value		.549	.098	.709	.184
Activation	Correlation			0.172	0.257	-0.499*
	P Value		-	.382	.187	.007
Monitoring	Correlation			-	-0.217	-0.046
	P Value				.267	.815
Non-verbal Inhibition	Correlation				-	-0.244
	P Value					.211

Overall, a significant negative correlation was observed between *verbal inhibitory control* and *activation*. The relationship between these measures is presented in Figure 4.3. As is evident from the figure, this correlation appears to be driven by just one patient, SO, who presented with a left parietal glioblastoma. Specifically, SO was significantly impaired on our aggregated activation measure, yet showed intact performance on the Stroop task. Indeed, when SO was removed from the correlational analysis, the negative correlation was no longer significant (R= -0.282, p= .162). As discussed in Chapter 6, SO's poor performance activating a response set is likely driven by residual effects of his chemotherapy treatment.
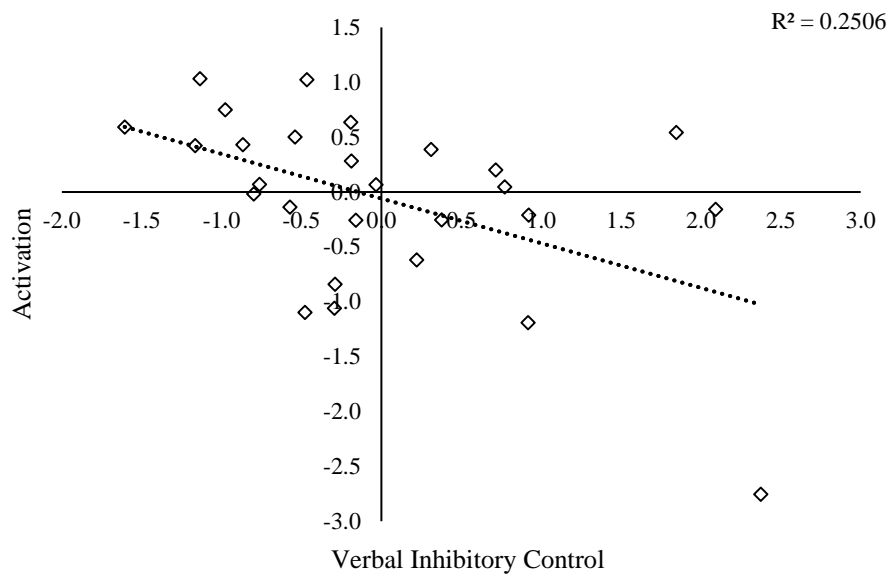
Figure 4.3. Scatterplot for the relationship between the aggregated measures of verbal inhibitory control and activation. The x axis represents the former measure, whilst the y axis represents the latter measure. As is evident, one patient appeared to be driving this effect, which is visible in the bottom right-hand corner.

Finally, it is worth noting that each of the above analyses assume that the specific measures we selected to contribute to each of our aggregate skill scores do indeed index a single common underlying cognitive skill. The next step in this research would be to test this hypothesis by examining whether the individual measures that contribute to the same core skill correlate with one another. It is difficult to test this prediction with a sample of only 28 participants. However, we do present some exploratory analysis and discussion of this in Appendix F.

Effects of tumour localisation. In order to ascertain whether there was any association between specific aggregated skill scores and lesion localisation, patients were first categorised into one of the following five groups using the methods outlined previously: left frontal (n=5), right frontal (n=8), left temporal (n=4), left parietal (n=6) and right posterior (n=4)[17]. To test each of our specific hypotheses, we first performed an omnibus one-way ANOVA using SPSS Software, considering all groups simultaneously, and then if there was a significant overall group effect, we specifically compared scores for the group of interest with those for the other groups considered together. Based on findings from previous studies, we

---

[17] One patient, AE was excluded from this analysis due to the presence of multiple lesions

made the following predictions regarding each of our core skill measures: 1) patients with frontal lesions would score significantly poorer than the remaining patients on measures of *activation*; 2) patients with right frontal lesions would score significantly poorer than the remaining patients on measures of *attentional maintenance* and *performance monitoring*; and 3) patients with left frontal lesions would score significantly poorer than the remaining patients on measures of *verbal inhibition*, whilst those with right frontal lesions would score significantly poorer than the remaining patients on measures of *non-verbal inhibition*.

However, there was no significant effect of group on the omnibus analysis for any of the five aggregated measures, using a p value of .05. Therefore, in no instance did we proceed to do more specific comparisons amongst subgroups.

Each of these non-significant findings are perhaps not surprising given the small number of patients in each localisation group, and the low levels of overlap between each of our tumour patients within each subgroup (see Chapter 2, Figure 2.2).

Finally, we also performed statistical power maps for each of the five aggregate skill scores, to ascertain whether it was possible to proceed with voxel-based analysis. Power maps specify whether there is sufficient statistical probability ($p < .05$) to detect damage to voxels in different neural regions. Here, our five aggregate skill scores were based on continuous data, and therefore we used the Wilcoxon-Mann-Whitney probability. However, there was insufficient power to detect a significant effect in any brain region at the familywise $p < .05$ false discovery rate, which was the case for all five measures. Therefore, in no instance did we proceed with more sophisticated voxel-based lesion analysis.

Discussion

The overarching aim of the current study was to isolate and identify the specific processes that contribute to higher-level cognitive control and their inter-relationships, using tasks that are as "pure" as possible. With this objective in mind, our starting point was the framework of Stuss and colleagues (1995; 2005), which proposes three major processes: *energisation, inhibition,* and *performance monitoring.* We further hypothesised that the process of energisation could be further defined into separable components based upon the *activation* of the schemata, and the *maintenance* and *sustained* activation of schemata over time. The major goals of this Chapter were two-fold. First, we aimed to investigate the inter-relationships between each of our putative measures in order to address more specific questions about the functional organisation of cognitive control. Second, we aimed to assess the anatomical regions that may be implicated in each of our putative measures of cognitive control. Each of these goals will be discussed below.

Investigating the inter-relationships between our cognitive skills

To investigate these putative control skills, we identified specific indices for each of our hypothesised cognitive functions, which were as selective as possible for the function under consideration. Elucidating measures to this level was a considerable advantage of the current study as it allowed us to make finer-grained inferences regarding processes of cognitive control in isolation from one another. Based on theories of regional specialisation within the PFC, we predicted that we would find evidence of fractionation between each of our core cognitive skills.

To summarise the overall findings, no significant correlations were observed between any of our aggregated measures, each of which were based on theoretical constructs of the cognitive processes they were designed to measure. Such findings are consistent with regional specialisation models of cognitive control, which assert that different types of control processes operate in a functionally distinct manner and are supported by different anatomical structures (see for example Stuss et al., 1995; 2005). However, our study sample was small, and we would need to replicate this finding on a larger sample before it could be considered secure. Further, on the basis of this study alone, we were unable to demonstrate that the different hypothesised processes were associated with distinctly different anatomical structures. Importantly, a larger sample size would be needed to confirm these findings, however until then, these null effects should be interpreted with caution.

Considering each of our specific hypotheses in turn, we hypothesised that the process of *energisation* proposed by Stuss and colleagues (1995; 2005)*,* could be further decomposed into separable, and functionally distinct subcomponents of *activation* and *attentional maintenance (sustained attention).* Indeed, separable anatomical regions have been implicated in the capacity to initiate (activate) schemata, and sustain that response (attentional maintenance), namely the superior medial regions of the PFC and right lateral PFC regions respectively (Alexander et al., 2005; Rueckert & Grafman, 1996; Shallice et al., 2008; Stuss et al., 1998; 2001; 2005; Wilkins et al., 1987). The non-significant association between our aggregated measures of *activation* and *attentional maintenance* is consistent with this hypothesis.

The absence of significant correlation between *performance monitoring* and *inhibitory control* measures is consistent with recent proposals that there may be two qualitatively different forms of conflict monitoring: one which operates in an anticipatory manner, to ensure errors never occur (called *pre* response conflict), and one which operates retrospectively to adjust performance parameters when this process fails and an error does occur (called *post* response conflict; see Ridderinkhof et al., 2004; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004b). Additional support for this proposal comes from neurophysiological evidence, with event-related potential (ERP) studies finding that *pre* and *post*-response conflict may be associated with differential patterns of timing, namely the N200 and error-related negativity (ERN) respectively (see Ridderinkhof et al., 2004b).

It is worth pointing out that some of the measures that contributed to our performance monitoring skill score were error measures (e.g., rates of inappropriate and perseverative errors on verbal and design fluency tasks contributed to this measure). It could be argued that these types of error measures index inhibitory control, rather than performance monitoring per se. For example, it has been proposed that perseverative errors might arise because a previously produced response becomes strongly reactivated during a subsequent trial, so that this becomes a type of "prepotent" response, that is difficult to inhibit (e.g., Papagno & Basso, 1996; Santo Pietro & Rigrodsky, 1986; Yamadori, 1981). Indeed, some have suggested that these previously active cognitive representations continue to persist for some time (e.g., Cohen & Dehaene, 1998), and may become amplified with each successive perseverative error. However, the absence of significant correlation between our aggregated measures of *performance monitoring* and *inhibitory control* suggests this is unlikely to be the case. It appears that perseverative (and task-inappropriate) errors may indeed constitute a

separate class of errors; however, of course, we would need to test this hypothesis on a larger sample before we could draw any firm conclusions.

Limitations of our tasks and measures

Considering the results as a whole, the absence of significant correlations amongst our four key measures may reflect artefacts of the way that our components were operationalised. For example, whilst there are several ways to operationalise post-error slowing, we opted to operationalise this construct using measures that are commonly used in the cognitive literature (the absolute mean difference in response latency between *post-error* trials and *post-correct* trials; e.g. Fellows & Farah, 2005; Kerns et al., 2005; see also Dutilh, van Ravenzwaaij, Nieuwenhuis, van der Maas, Forstmann, & Wagenmakers, 2012). Recently however, the appropriateness of this approach has been called into question. According to Dutilh and colleagues (2012), despite the strong face validity of this conventional method, it may in fact be confounded by global changes in *motivation* and *response caution*. For example, on the initial trials, a participant's motivation and ability are high, and they are likely to adopt a cautious response style. However, during the latter half of the task, the individual may be less motivated and cautious, and thus commit disproportionately more errors. Accordingly, much of the post-error *slowing* measure will originate from the second half of the trials, whilst much of the post-*correct* latency measures will originate from the first half when responses are faster and more accurate. Dutilh and colleagues (2012) argue that this comparison may result in spurious effects, that is, either an artificial PES, or a failure to detect genuine effects. Whilst this is an important consideration, close inspection of our data reveal a non-specific pattern of accuracy performance. That is, the majority of patients appeared to commit equally as many errors during the first half of the trials as they did on the second half. Therefore, this limitation may not be as critical as it first appears, but is nonetheless important to consider for future studies that use this paradigm.

On a similar note, the validity of post-error slowing has been called into question as a measure of performance monitoring. For example, it has been proposed that post-error slowing arises due to a persistent deficit that caused the initial error (Gehring, Goss, Coles, Meyer & Donchin, 1993). According to this account, individuals who exhibit difficulty on any component or construct will also have difficulty adjusting their performance following the commission of an error – that is, they will show global impairments, regardless of cognitive modality or domain. Whilst we acknowledge that it is certainly important to establish the validity of our measures, the absence of significant associations between

performance monitoring and the remaining core skills appears inconsistent with this account. Importantly though, a larger sample size would be needed to confirm this result.

A further artefact of our tasks may be due to the way that our error rates were measured and defined. For each of the verbal and non-verbal fluency conditions, errors were calculated as a proportion of the total number of items produced. We reasoned that this approach may control for differences in the opportunities for error between those who generate many responses and those whose responses are sparse. However, our proportional measure is not without its limitations. Specifically, our error proportion measures may also be influenced by overall task accuracy: for example, those who are poor on the task overall may produce greater numbers of errors *of all kinds*, including the specific error types we focused on here. If this is the case, our error measure may not constitute an *independent* measure of monitoring; it may be contaminated by other capacities, such as verbal (or nonverbal) inhibitory control. To overcome this confound, future studies may wish to use raw error rates and factor out other, non-related factors, such as activation, using multiple regression. This was not possible in the current study due to the small sample size. Alternatively, performance monitoring could be assessed using electrophysiological markers, such as event-related potentials.

It is also worth noting that each of the error measures that contributed to performance monitoring was based on the participant committing an error. However, such occurrences were generally rare in our sample, and thus our aggregated measure likely had a skewed distribution that may not be representative of neurological tumour patients, nor reflect the true nature of performance monitoring. In Appendix F, we briefly explore some aspects of the validity of our various performance monitoring measures (although these analyses would need to be confirmed in a larger sample).

We turn now to potential artefacts and limitations within our measures of inhibitory control. It is worth noting that within the BLAST subtest of the Stroop task, the proportion of correct responses on the incongruent condition was directly compared to that of the congruent condition in order to derive a measure of verbal inhibitory control. However, this form of comparison may have created artefacts not necessarily attributable to inhibitory control deficits. Specifically, in the "baseline" congruent condition, participants are directly provided with the word name itself. This presentation may create facilitatory effects, wherein words are retrieved faster from the lexicon as a result of the word stimulus provided, particularly for those with lexical selection difficulties. Accordingly, any differences in response latencies

between congruent and incongruent items may reflect the facilitatory effect of being provided with the target colour name rather than the inhibitory effect of viewing a conflicting colour name. To overcome this potential artefact, future studies should endeavour to include a neutral condition in the Stroop task which may tease apart genuine effects of interference from those of lexical assess (see McLeod, 1991 for discussion).

A potential limitation of our non-verbal inhibitory control measure concerns the tasks that contributed to its operationalisation: namely, design fluency and the Choice and Prepare RT tasks. Specifically, these tasks create a "prepotent" response by incorporating frequent repeats of specific trials. This is a considerably weaker method when compared to methods that capitalise on "naturally" prepotent responses (such as reading a word name as opposed to identifying the word's ink colour). These trial conditions were selected for the following reasons: 1) they were consistent with previous studies; and 2) it was important to mitigate any confounds of fatigue, concentration, and motivation by keeping trials to a minimum. Despite these justifications, it is entirely possible that these measures are weak at indexing inhibitory control *per se*, and thus, any results should be interpreted cautiously. Future studies should endeavour to include more trials in order to draw more definitive conclusions, whilst also validating these measures against alternative ones hypothesised to index the same processes. To this end, we present exploratory correlations between each of our individual measures in Appendix F, however, a larger sample size would be needed to confirm any findings.

Moving now to our aggregated measure of activation, it is worth noting that performance on this skill was derived entirely from tasks that required motor responses. Whilst this is a common method to measure energisation/ activation (e.g. Stuss et al., 2005; Alexander et al., 2005), it may nonetheless be confounded by motor difficulties and/or general speed problems. This is a potentially important limitation as a number of our patients had recently undergone chemotherapy and/or radiation therapy, of which may produce generalised cognitive and motor slowing (Douw et al., 2009; Schatz, Kramer, Ablin, & Matthay, 2000). To overcome such confounds, future studies should endeavour to include a range of response types from which to compare, including motor, verbal, and eye-movements, as well as comparing and contrasting performance of those who receive chemotherapy and/or radiation to those who do not.

Further, we used an aggregated measure of activation as a marker of energisation. We predicted, based on previous findings, that this process would be localised to frontal regions (see Alexander et al., 2005; Stuss et al., 1995; 2005). However, the non-selective

impairments across both frontal and posterior patients (see Appendix F) suggests that whilst there may indeed be such a process, speed itself might be an outcome that reflects the efficacy of multiple processes. Accordingly, it is likely that activation is affected by global factors, and may not be specific enough as a measure of cognitive control. To further explore this issue, future studies should first aim to validate the individual measures of activation. Whilst we present some exploratory data of this in Appendix F, our small sample size limits any definitive conclusions.

Effects of lesion location on our core cognitive processes.

Based on previous observations of regional specialisation within the PFC (see especially Stuss et al., 2005), we made the following predictions about the anatomical correlates of each of our core cognitive processes. First, given previous findings of right frontal PFC involvement in *performance monitoring*, we predicted that patients with right frontal lesions would exhibit significantly poorer performance on our performance monitoring measure than the remaining groups. Second, we predicted that patients with left or right frontal lesions would perform significantly more poorly on our aggregated measure of *activation* than patients in the remaining tumour groups. Third, we predicted that patients with right frontal lesions would perform significantly more poorly on our measure of *attentional maintenance* than the remaining patients. Finally, we predicted that patients with frontal lesions would perform significantly more poorly on *inhibition* compared to the remaining tumour groups. Specifically, we predicted that those with left frontal lesions would perform significantly more poorly on our *verbal* measure of inhibition, whilst those with right frontal lesions would perform significantly more poorly on our *non-verbal* measure of inhibition, relative to the remaining tumour groups. Despite these predictions, no significant effects of localisation were observed. Below, we discuss potential reasons and implications of this.

An important issue to consider when evaluating these results is the characteristics of our patient sample itself. Our sample included 28 undifferentiated patients with lesions to both anterior and posterior regions. Further, as shown in Chapter 2, Figure 2.2., few patients had lesions extending into the superior medial and right PFC regions. In contrast, the 38-case sample of Stuss and colleagues (2005) consisted entirely of those with focal frontal lesions, and there was much greater lesion overlap in crucial areas necessary for cognitive control (e.g., superior medial PFC, inferior medial PFC, left lateral PFC, and right lateral PFC). Importantly, Stuss and colleagues (2005) never actually established that their selected

patterns of performance were indeed specifically associated with *frontal* – rather than posterior – lesions. Our results suggest that this is by no means a certainty.

Further, as is discussed below (General Discussion, Chapter 6), lesion localisation was widely dispersed in the current study, and very few patients had overlapping lesions in any particular brain region (Chapter 2, see Figure 2.2). Consequently, we were restricted to comparisons across broadly defined lesion groups. However, this approach is far from ideal, because only a small proportion of patients in each localisation group may have a lesion in the specific region hypothesised to be implicated in the function of interest. As a result, genuine associations between location and cognitive performance may be washed out by the grouping process. To overcome such difficulties, future studies should endeavour to include a larger sample size, which would enable for more sophisticated methods to infer brain-behaviour relationships, such as voxel-based lesion symptom mapping.

Related to this, it is likely that the absence of significant lesion effects reflect the patient sample itself. Indeed, it is entirely possible that we did not have representative cases in our sample, wherein there was no coverage of the regions and their associated deficits that we were hypothesising for. The representativeness of the sample is always a challenge in the neuropsychological domain, however, had the sample included such cases, it is likely that we would find substantially different results.

Whilst our data demonstrated the overall sensitivity of the core skills approach to measuring cognitive control, it did not appear to be very specific, at least with respect to the localisation of the lesion in patients with significant impairments in a particular skill. Indeed, despite being conceptualised as "frontal" processes, a high proportion of patients with posterior lesions exhibited significant impairment across our skill measures (see especially Appendix F). Such observations may suggest that our measures reflect generic cognitive functions rather than processes isolated to PFC regions. This is an issue that appears to have been side-stepped in the neuropsychological literature, wherein the vast majority of studies that examine "frontal" processes have exclusively included patients *only* with frontal lesions. Indeed, a number of studies that have included a broader sample of patients have found that traditional 'frontal" tasks, such as the Wisconsin Card Sorting Test (WCST), fails to discriminate between frontal and non-frontal patients (see Goldstein, Obrzut, John, Ledakis, & Armstrong, 2004; Nyhus & Barceló, 2009). For example, lesions to non-frontal regions, such as temporal, subcortical, hippocampal, and cerebellar have been implicated in impaired performance on the WCST (see Nyhus & Barceló, 2009).

One problem here is that many tasks designed to measure "frontal" processes are difficult in a number of diverse ways. Consequently, they are likely to place greater demands on a whole host of resources and cognitive processes, including those not commonly associated with PFC functions. Accordingly, it may be the case that before we can adequately investigate brain-behaviour relationships with respect to "frontal" tasks, we need a better definition of "control" processes that focuses more specifically on those associated with the prefrontal cortex.

Finally, it is worth bearing in mind that many studies of patients with frontal damage report marked dissociations between their performance on laboratory tasks and their behaviour in everyday contexts (see Chaytor et al., 2006). Indeed, it is conceivable that certain impairments are particularly manifested in uncontrolled and unconstrained environments, wherein multiple stimuli must be attended to, integrated, monitored, maintained, and responded to, within a short period of time (Rueckert & Grafman, 1996). The tasks examined here index only certain types of processes under very constrained conditions. Accordingly, future studies may wish to compare performance on very simple, core component measures, such as those utilised in the current study, to performance on challenging tasks of daily living.

Chapter 5: The Relationship between Language and Cognitive Control

Introduction

One dominant assumption within the literature is that there are language-specific cognitive processes, which are only used within the language domain. Traditionally, researchers studying language functions have made a distinction between what they consider to be "core" language functions (such as auditory word recognition, word retrieval, and articulatory-motor planning) and other more general cognitive functions that are not strictly language processes, but which support performance (e.g. attention, inhibitory control). However, recent research suggests that some of these impairments affecting these "extra-linguistic" processes can be highly specific to verbal materials. For example, Hamilton and Martin (2005) reported that their patient with non-fluent aphasia, who performed poorly on the incongruent condition of the Stroop task, was not similarly impaired on a nonverbal task requiring inhibitory control – namely, the antisaccade task. That is, there appears to be a set of control functions that are highly specific to language. Clearly, the relationship between language and cognitive control is more complex than previously believed, and is worthy of further investigation; indeed, the evaluation of this relationship can further expand our theoretical understanding of the functional architecture of language, and the mechanisms that underpin linguistic processes. Such insights may ultimately help us to gain a better understanding of how language breaks down after brain damage, which in turn can inform assessment and treatment guidelines in neurological populations. Consequently, the final theoretical objective of this thesis is to address the relationship between selected language skills and more generalised aspects of cognitive functioning.

There is now a large body of evidence indicating that the neural structures supporting language function extend considerably beyond the classical language areas of Broca's and Wernicke's regions. Indeed, Vigneau and Colleagues (2006) conducted a large-scale meta-analysis of the neural regions activated during various types of language tasks in healthy, unimpaired individuals. Their meta-analysis involved 129 published fMRI studies and excluded occipital regions, which might be implicated in tasks that require visual processing. Overall, they reported the cortical sites associated with language processes extended well beyond the classical anatomical boundaries of Broca's and Wernicke's area (Vigneau et al., 2006). For example, a diverse range of frontal and temporal clusters were identified for phonological processing, including the dorsal part of the pars triangularis of the inferior frontal gyrus, anterior and middle regions of the superior temporal sulcus, and the middle

temporal sulcus. If we extend our consideration to discourse comprehension, additional regions are likely to be implicated (e.g. orbitofrontal cortex and anterior temporal lobes; see Barbey, Colom, & Grafman, 2013).

Moreover, evidence from electrical stimulation studies indicates that many of these regions are *critical* for effective language performance. Using cortical stimulation mapping, Sanai and colleagues (2008) assessed the performance of 250 patients undergoing tumour resection surgery on a number of simple language tasks, such as picture naming, and word reading (see chapter 2 for further details). They found that in 92 of the 151 patients with frontal lesions, stimulation of at least one region in the frontal cortex resulted in a specific language impairment, such as a temporary loss of the ability to produce words (*speech arrest*), difficulty recalling names and words (*anomia*), and a difficulty understanding written words (*alexia*; Sanai et al., 2008).

Importantly, the above findings argue against a modular organisation of language often characterised by the classical model, and instead suggest that a widespread anatomical network is implicated in linguistic functions. Such observations further support the assertion that conventional aphasia models fail to adequately capture the true nature of language impairments. These findings support the idea that the recruitment of a wide range of cognitive processes are necessary to perform basic elements of language, such as single-word tasks, as well as more complex operations involving sentence-level comprehension and production (Alexander, 2006; Hula & McNeil, 2008; Murray, 2012; Shao et al., 2014). Importantly however, despite this converging evidence, there remains little research investigating whether regions extending beyond the traditional language regions represent genuine language-specific areas, or instead perform less language-specific functions, such as general attention or inhibitory control.

Specifically, many of the cognitive functions proposed to play a role in linguistic processes fall into the general category of cognitive control skills. As mentioned in Chapter 1, cognitive control refers to a collection of skills mediated by the frontal regions that operate under top-down control. Such skills are deemed necessary for goal-orientated behaviour that acts to guide actions in accordance with current task demands (Badre, 2008; Badre et al., 2009; Novick et al., 2009). To demonstrate this, consider the Stroop task, described in previous Chapters. Successful completion of this task requires formulating a response goal (saying the colour name rather than the word name), maintaining that goal in an activated state throughout the task, and of course, using that goal to guide behaviour. Specifically, it

remains unclear the extent to which some of these skills operate in a generalised manner over all kinds of domains and materials equally, or whether there are specific sets of cognitive control serving verbal and nonverbal domains. Research addressing this question has led to conflicting accounts and has yet to be answered definitively (Tsuchida & Fellows, 2013).

The subsequent sections will focus on two cognitive control skills held to play a fundamental role in goal-driven behaviour and linguistic functions: *Inhibitory control* and *Internally-driven generation*. Particular emphasis will be on the role they play in language, both at the single-word and sentence-level. From here, the current study will attempt to address the question of whether these two cognitive skills operate in a domain general or domain specific manner.

Domain Specific versus Domain General

Inhibitory Control and Conflict Resolution

As discussed in Chapter 1, a core component of cognitive control is the ability to select the representation or action plan that best suits our current goals whilst suppressing those that may be concurrently activated (Badre et al., 2009; Hamilton & Martin, 2005; Novick et al., 2009). This ability is crucial for everyday linguistic functioning. For example, during a conversation, multiple conceptual and lexical representations are likely to be simultaneously activated. In such situations, effective production and comprehension performance dictates that only the most relevant material is attended to; representations not directly relevant to the task-goal must be actively suppressed.

To illustrate this point, again consider the Stroop task. As described in previous Chapters, the Stroop task requires participants to override the habitual, most overlearned response to words – which is to read them – and instead favour the less common alternative, of naming the colour they are presented in. Failure on this task has been conceptualised as reflecting difficulties with either language processes, or general inhibition/ conflict resolution (e.g. Novick et al., 2010; Hamilton and Martin, 2005). Importantly however, recent views suggest that poor inhibitory control/conflict resolution may be a language difficulty in itself. Indeed, a number of patient and imaging studies have reported that the Stroop task is associated with the left inferior frontal gyrus (LIFG) – an area held to support top-down control processes in a range of language tasks that appear to require the resolution of conflict between verbal representations and the suppression of irrelevant ones (Thompson-Schill, D'Esposito, Aguirre & Farah, 1997; Thompson-Schill et al., 1998; Novick et al., 2010).

Patients with damage to the LIFG who perform poorly on the Stroop task are often also impaired on other tasks that might arguably involve similar skills. For example, they may be impaired on more open-ended tasks, such as verbal fluency and verb generation (e.g. Novick et al., 2009; Robinson et al., 2005; Thompson-Schill et al., 1997). As discussed previously, such generation tasks are unconstrained: the target items are not provided to the individual and instead they must utilise internal strategies to come up with items of their own. For example, Novick (2009) reported a case of an individual who had particular difficulty generating word items on a letter fluency task (e.g., generating words starting with "f"). This patient also had selective difficulty generating verbs from nouns where there was no single strongly preferred option (e.g. a person could respond to "*rope*" with "pull", "tie", "hang"). However, he displayed intact performance when the noun he was given was associated with one overwhelming preferred verb response (e.g. *nun* -> pray). Similar findings have been reported in a number of other patient studies involving damage to the LIFG (e.g. Novick et al., 2009; Robinson et al., 2005; Swick, Ashley, & Turken, 2008; Thompson-Schill et al., 1999; Thompson-Schill et al., 1997). Such observations are consistent with evidence from fMRI studies of healthy individuals, which have observed greater activation of the LIFG on verb generation tasks involving "weak items", when compared to those involving "strong" ones (Thompson-Schill et al., 1997).[18] Taken together, this evidence suggests that all three of these tasks – Stroop, letter fluency and verb generation – may rely upon a common set of cognitive control skills that help guide the selection of a verbal response and ensure that it wins the competition for selection over other concurrently activated items.

These types of impairments can be understood in terms of spreading activation theories of language (discussed in Chapter 1). These models propose that during the production of a word (for example, naming a picture), multiple lexical representations become available, in direct proportion to their semantic "fit" with the concept in mind. The node that receives the greatest activation is then selected for production (see Dell, 1986; Levelt, 1999). Within this spreading activation framework, difficulty arises under relatively unconstrained conditions, wherein a noun activates several appropriate, and therefore acceptable, lexical nodes simultaneously (Robinson et al., 2005; Thompson-Schill et al.,

---

[18] It has been suggested that the LIFG's involvement in the "weak" items on the verb generation tasks might not be specifically related to selection amongst multiple alternatives, but rather might reflect the challenges these items place on lexical search processes. According to this view, items such as "rope" are more difficult than "scissors" because they do not automatically activate a word associate. Therefore, the person must search through their lexion to find an appropriate verb (Martin & Chao, 2006; Wagner et al., 2014).

1997). In these circumstances, one response must be preferentially activated over other competing alternatives (Robinson et al. 2005); it is important to ensure that only the most relevant and appropriate lexical representation is selected. A number of mechanisms are thought to be dedicated to overseeing this flow of activation, which operate to either actively enhance goal-relevant representations and/or inhibit inappropriate ones (Biegler et al., 2008; Hamilton & Martin, 2005; January et al., 2009; Schnur, Schwartz, Brecher, & Hodgson, 2006; Thompson-Schill et al., 1997; Wilshire & McCarthy, 2002).

Taken together, these observations suggest that the need to oversee verbal response selection is required in almost all language settings (e.g. single and sentence-level, naming, and generation), even in tasks not traditionally associated with inhibitory control, such as verb generation and verbal fluency.

Also consistent with this account is the observation that many patients with this hypothesised deficit exhibit reduced spontaneous speech (sometimes non-fluent, sometimes just very sparse), yet relatively well-preserved language skills under more constrained conditions and external cues (see also Robinson, Blair & Cipolotti, 1998; Robinson et al., 2005). As discussed above, these difficulties at the sentence-level may also be understood in terms of a similar framework. Specifically, sentence-level speech production is distinct from single-word tasks, in that a message intention must be formulated in order to guide the overall sentence structure; this intention must remain activated and maintained to exert top-down control of the language system, so that task-relevant representations are selected and irrelevant ones are suppressed (Novick et al., 2010). Again, evidence suggests that the LIFG is crucial in this situation. For example, Robinson and colleagues (1998) describe the sentence production performance of one individual (A.N.G), who presented with a lesion impinging on the LIFG due to a malignant meningioma. Despite intact naming, reading and repetition, A.N.G. exhibited significant difficulty producing spontaneous speech (a speech pattern often referred to as *dynamic aphasia*; Robinson, Blair, & Cipolotti, 1998). In one task, these authors ask A.N.G. to generate a sentence based on a simple prompt word. When the prompt word was a common one (e.g., 'table') the patient performed more poorly than when the prompt word was a less common one (e.g., 'Ghandi'). Robinson and colleagues (1998) argued that the common prompt words placed greater demands on a conflict resolution mechanism, because they are consistent with a wide number of different sentence types (e.g., *table* may elicit sentences to do with sitting, wiping, building, polishing, whereas Ghandi is likely to elicit sentences related to Ghandi's fame). In another task, A.N.G. was given a

sentence fragment and asked to complete it by adding a final word. A.N.G. performed more poorly on sentences that were associated with many possible completions (e.g. '*Helen reached up to dust the…*') than on those that offered one single predominant completion (e.g. '*Dogs have a good sense of…*'). Again, this pattern is also consistent with a failure at the level of top down lexical control (see Robinson et al., 2005, for a subsequent report of a similar case; see also Novick et al., 2009 for a similar case).

Is top down verbal control a "language" function?

So far we have demonstrated that there may be a type of top-down control that is necessary for linguistic processing. However, it currently remains unclear whether this type of control is specific to verbal performance or instead operates across all modalities and information types (Novick et al., 2009; Robinson et al., 2005).

Much of the evidence from healthy participants supports the view that there are core domain-general executive control functions that are closely interrelated. For example, using factor analysis to examine interrelationship between various executive function skills in neurologically healthy participants, Miyake and colleagues (2000) found both the verbal and nonverbal analogues of inhibition loaded onto a single factor, which they referred to as the inhibition factor (Miyake, Friedman, Emerson, Witzki, Howerter, & Wagner, 2000). Further, in a later study that examined the relations between different types of inhibition, Friedman and Miyake (2004) found that performance on the conventional verbal Stroop task and performance on a nonverbal inhibitory control task - the antisaccade task - was highly correlated in a sample of 220 undergraduate students. Taken together, these findings do appear to suggest an underlying commonality between the inhibitory processes called upon to resolve conflict in different domains.

In contrast to the above findings, several studies of brain-damaged patients have found evidence that challenges the notion of a unified control mechanism, and instead argue toward a domain-specific account of top-down verbal control. Specifically, in an earlier study, Hamilton and Martin (2005) assessed a single patient (M.L) who presented with damage to the LIFG following a cerebral vascular accident. Both the verbal Stroop and antisaccade task were administered, each of which were used in the factor analysis of healthy individuals described above (Miyake et al., 2000). The antisaccade task involved a series of arrows pointing to either the left or right-hand side of a computer screen, with M.L required to press a key corresponding to each arrow direction. Similar to the verbal Stroop paradigm,

the task contained congruent trials (left-pointing arrow on left side of screen), neutral trials (arrows presented in the middle of the screen), and incongruent trials (left-pointing arrow on right side of screen) that called for the suppression of a prepotent response to successfully achieve the task. Overall, M.L exhibited exaggerated interference effects on the Stroop task, as evidenced by longer response times (RT's) on the incongruent items, however, M.L's performance was comparable to controls on the antisaccade task. Importantly, this finding could not be attributed to the antisaccade task being generally easier, given controls performed more poorly on it relative to the Stroop task (Hamilton & Martin, 2005). M.L's performance pattern argues against a unified inhibition theory described by Miyake and Colleagues (2000), and instead supports the notion that verbal top-down control can be further decomposed into finer grained, modality-specific functions that support different types of material.

More recently, Geddes' and colleagues (2014) evaluated material-specific processing in a sample of eight patients with lesions to the left or right ventro-lateral PFC (VLPFC). All patients were assessed at least one-year post-injury (mean = 3.3 years, range = 1.2 -6.4 years) on verbal and spatial measures of interference. Verbal inference was assessed using the traditional Stroop paradigm. As described previously, this task requires participants to overcome their dominant, or "prepotent" response of reading the word and instead must name the colour that the word is written in. In contrast, spatial interference was assessed using a speeded version of the Eriksen flanker task. This paradigm requires participants to view a series of five arrows and indicate the direction of the central arrow as quickly as possible. There are two conditions: congruent ($< < < < <$ or $> > > > >$) and incongruent ($< < > < <$ or $> > < > >$) (see Geddes, Tsuchida, Ashley, Swick, & Fellows, 2014). Overall, patients exhibited evidence of a double dissociation, with lesions to the left VLPFC associated with exaggerated interference effects on the Stroop task, but not on the Eriksen flanker task. Conversely, lesions to the right VLPFC were associated with the opposite pattern of performance: exaggerated interference effects on the Eriksen flanker task, but not on the Stroop. Such results led Geddes' and colleagues (2014) to conclude that cognitive control may operate in a lateralised, domain-specific manner.

Of course, both these studies focused on relatively small cases, and only the former directly assessed verbal and non-verbal functions. Therefore, to address the generalisability of these findings, it is important to expand on the above studies with a larger sample size. Further, it is also necessary to examine cognitive and language functions more generally in

order to isolate inhibitory effects from other, more generalised cognitive deficits. Finally, it is important to include an undifferentiated sample of both anterior and posterior cases in order to explore the validity of these "frontal" tasks.

Internally-driven generation – domain general or domain specific?

Extending our consideration beyond tasks that explicitly incorporate interference, we now consider other tasks that require more general internally-driven generation and selection of actions. The generation of actions can be broadly classified along a continuum from externally to internally-specified; an externally specified action is largely automatic and is cued directly by an environmental trigger that specifies the appropriate action to take (Tremblay & Gracco, 2006). Conversely, an internally specified action is cued by an internal event, and requires a more conscious and controlled decision regarding the appropriate action to perform, and when to initiate it. Tasks that exemplify this type of action selection are those where a spontaneous response must be initiated by the individual without any direct cues from the environment (Tremblay & Gracco, 2006). A common example used in language research is the verbal fluency task. There is also a nonverbal analogue of this task, called the design fluency task, which involves generating novel geometric designs (Robinson, Shallice, Bozzali & Cipolotti, 2012). By comparing patients' performance on these two tasks, we can gain insight into whether any problems exhibited in the verbal domain extend to tasks not involving verbal materials. Such insights can ultimately provide a new source of knowledge regarding the nature of these processes, that is, whether they operate in a domain general or domain specific manner.

Similar to the Stroop literature, some studies have examined this question using healthy non-brain-damaged populations. In a recent study, Suchy and Colleagues (2010) assessed 61 neurologically healthy older control participants on a range of fluency tasks. These included standard letter and category fluency tasks, a design fluency task (with no switching requirement), and also a motor sequence fluency task, where individuals were required to produce as many unique sequences of three hand movements as possible in an allocated time period. Motor sequence fluency was found to be a significant predictor of design fluency, however performance on verbal fluency was not a significant predictor of either (Suchy, Kraybill, Gidley, & Larson, 2010). These results challenge a unitary construct of fluency, and instead suggest that verbal and nonverbal indices of internally-driven generation are functionally distinct and subserved by different systems. Importantly however, only 30-40% of the overall variance was accounted for in these models, thus any

interpretations should be tentative. This highlights the need for further investigation regarding the domain specificity of these processes in a neurological population.

Turning now to studies of brain-damaged patients, some of which suggest that all fluency tasks might share some common cognitive elements. For example, Butler and colleagues (1993) assessed 17 patients with frontal neurological tumours on a series of fluency tasks, including letter fluency, design fluency (free and fixed conditions; see below), category fluency (e.g. possible jobs), and other more complex fluency tasks (e.g. alternate and new uses for objects; Butler, Rorsman, Hill, & Tuma, 1993). Importantly, patients with left frontal tumours scored more poorly on all verbal and non-verbal tasks, although the differences did not always reach significance.

However, in contrast to these findings, many patient studies, and studies of healthy controls, suggest there may be differential involvement of the left and right hemisphere depending on the nature of the task materials. Specifically, left-hemispheric lesions have generally been associated with deficits on verbal fluency tasks, whilst right hemispheric lesions have often been associated with deficits on nonverbal fluency tasks (Jones-Gotman & Milner, 1977; Lee, Strauss, McCloskey, Loring, & Drane, 1996; Miceli, Caltagirone, Gainotti, Masullo, & Silveri, 1981; Robinson et al., 2012; Ruff, Allen, Farrow, Niemann, & Wylie, 1994; Stuss, et al., 1998). Importantly though, not all studies support this idea of material-specific processing and hemispheric laterality; some have reported disrupted non-verbal fluency following left-hemispheric lesions (Jones-Gotman & Milner, 1977), and somewhat diminished verbal fluency following right-hemispheric lesions (Bruyer & Tuyumbu, 1980; Butler et al., 1993; Martin, Loring, Meador, & Lee, 1990; Miceli et al., 1981). For example, in one notable study, verbal and design fluency performance (measured using the D-KEFS) was assessed in an undifferentiated sample of 11 patients with focal frontal lesions. Patients were assessed on average 9.9 years following their injury onset and all had lesions confined to ventral and dorsolateral PFC (left hemisphere: 6, right hemisphere: 5). Overall, patients showed significant impairment - defined as fewer responses/designs produced - on both the verbal and non-verbal fluency tasks, when compared to neurologically healthy controls. Specifically, patients with left frontal lesions showed significantly more impairment on the verbal fluency task, but both left and right-hemispheric patients performed comparably on the design fluency task (Baldo, Shimamura, Delis, Kramer, & Kaplan, 2001). This finding suggests that successful performance on design fluency tasks involves bilateral

involvement of the PFC, whilst performance on verbal fluency tasks may place greater demands on left-hemispheric regions.

Dissociations between verbal and nonverbal fluency tasks have also been observed at the single case level. For example, Robinson and colleagues (2005) assessed a patient, C.H., who presented with a non-fluent form of aphasia characterised by sparse, but well-formed, spontaneous speech (dynamic aphasia). C.H. presented with a lesion localised to the LIFG. Overall, C.H. displayed severely reduced spontaneous speech output, characterised by low speech rate and poor initiation of conversation. However, C.H.'s performance was intact on more constrained tasks, such as repetition, picture naming, word comprehension (Robinson et al., 2005). These results led Robinson and colleagues (2005) to suggest that C.H.'s language difficulties were caused by a failure of cognitive control. Importantly though, this single case study does not enable us to address the question as to whether such problems are exclusive to patients with left hemispheric lesions, or whether they would be equally common in those with lesions to the right hemisphere.

Robinson and colleagues (2012) recently addressed this question more systematically in a group of seventy-two patients with focal frontal and posterior lesions (47 frontal; 20 posterior: 12 left and 8 right). The patients completed a range of verbal and nonverbal fluency tasks including; i) Verbal Fluency, consisting of a category and letter condition; ii) Design Fluency that utilised two conditions; a free condition that required individuals to generate as many designs as possible within five minutes, which did not represent real objects, nor were derived from them; and a fixed condition that had identical demands, except each design had to consist of four lines which could be straight or curved; iii) Gesture Fluency, wherein participants must generate as many upper-limb movements as possible in two conditions: *meaningful* and *meaningless*; and iv) Ideational Fluency, wherein participants must generate as many possible uses of an object in 90 seconds. Overall, patients with frontal lesions exhibited impairment on all fluency tasks, relative to healthy control participants; specifically, there were significant correlations between frontal patients' performance on the letter fluency task, the category fluency task, the Gesture Fluency task (meaningful gestures condition), and the ideational fluency task. However, no significant correlations were observed between the letter fluency and design fluency tasks. Interestingly, when compared to patients with posterior lesions, only design fluency and letter fluency were selectively impaired in the frontal patients. That is, all fluency tasks were sensitive to frontal damage, yet only letter and design fluency showed frontal specificity. Further, there appeared to be some

lateralisation, with left frontal patients most impaired on letter fluency compared to right frontal patients; conversely, right frontal patients were more impaired on the design fluency task. These results led Robinson and colleagues (2012) to conclude that letter fluency and design fluency are likely underpinned by different cortical substrates, which provides support for a domain-specific view of internally-driven initiation.

The Current Study: Task Selection and Hypotheses

In the current study, the overarching aim was to investigate whether crucial aspects of cognitive control operate within a domain-general, or domain-specific manner. To address this question, we selected measures that were as "pure" as possible, and which can be performed by a wide range of patients with different types of brain damage. Specifically, we begin by identifying tasks/measures that are commonly used to assess inhibition and/or internally-driven action generation. Following this, we will then discuss the specific predictions for each of our selected measures.

Inhibitory Control. As summarised above, inhibitory control is commonly assessed using tasks that involve inhibiting a well-learned, "prepotent" stimulus-response relationship so that another, less familiar or well learned response can be made. In the current study, to assess inhibitory control of verbal actions, we used the Stroop task. Conversely, to assess nonverbal inhibitory control, we used three tasks: 1) the antisaccade task; 2) the Choice and Prepare RT tasks; and 3) the "filter" condition of the Design Fluency task. The rationale for selecting each of these tasks/measures is set out in Chapter 4. Importantly, the Stroop and antisaccade task were selected due to their capacity to serve as verbal and non-verbal analogues of one another (see Hamilton & Martin, 2005). It is also worth noting that the two RT tasks – the Choice and prepare task – involve responding to letters, so they are not entirely "nonverbal" in content. However, the response they require is always nonverbal (a button press).

Turning now to the specific measures we derived from these tasks, we reasoned that on the Stroop task, impaired verbal inhibitory control might manifest itself in two ways. It might induce the person to produce disproportionately high rates of errors on incongruent items, some of which are likely to be substitutions of the "prepotent" word name response for the more appropriate, colour name response. Alternatively, impaired inhibitory control might manifest itself in abnormally prolonged response latencies and disproportionately reduced accuracy to the incongruent (high interference) items. For these reasons, in order to

operationalise verbal inhibition, we obtained two separate measures from this task: 1) a latency difference measure, expressed as the average latency on incongruent trial as a proportion of the average latency on congruent trials; and 2) an accuracy difference measure, which expresses the difference in the percentage of items correct on the incongruent and congruent conditions respectively.

To operationalise non-verbal inhibition, we used the following measures, all of which were explained and used in Chapter 4: 1) overall accuracy on the antisaccade task; 2) on the Choice and Prepare RT tasks, the difference in average latencies to the stimulus 'A' and average latencies to the remaining three stimuli (B,C, or D); and 3) on the design fluency task, the proportion of designs generated on the 'filter' condition, relative to the 'basic' condition.

Internally-driven action selection/generation. Tasks that have previously been used to operationalise internally-driven action selection/generation are those that require the free and spontaneous generation of actions according to some broader guiding principle, and where there are no external cues to fully constrain the final choice of action. Common examples are verbal fluency and design fluency tasks (Baldo et al., 2001; Robinson et al., 2012; Shao et al., 2014). In the current study, we adopt the same approach by comparing two tasks: verbal fluency and design fluency.

Specifically, we operationalised internally-driven generation of verbal responses using the following measures from the verbal fluency task: 1) total number of words produced on the letter fluency condition (FAS); and 2) total number of words produced on the category fluency condition. Based on a number of studies above (e.g. Baldo et al., 2001; Robinson et al., 2012), the first of these two measures would be predicted to be the most sensitive to impairments in internally-driven generation. We also included one further task that may be useful for assessing internally driven generation of verbal responses: the Verb Generation task. Specifically, an impairment involving this skill would be predicted to disproportionately affect the low response strength condition of this task, where the noun stimulus is compatible with several equally appropriate responses (e.g. rope -> *hang, tie, throw*), and thus results in a failure to activate an internally initiated response. Accordingly, we operationalised this pattern of performance in the same way as for the BLAST (see Chapter 2, see also Faulkner et al., 2017): we calculated and compared the average number of incorrect items on the low response strength condition, to the average number of incorrect items on the high response strength condition (e.g., scissors -> *cut*).

To operationalise nonverbal internally-driven generation, we used the following measures from the design fluency task: 1) total number of designs initiated on the *Basic* condition; 2) total number of designs initiated on the *Filter* condition; and 3) total number of designs initiated on the *Switch* condition.

Our predictions with respect to inhibitory control were as follows. Based on the hypothesis that inhibitory control is domain-specific (see Geddes et al., 2014; Hamilton & Martin, 2005), we predicted that there would be no association between the various measures of verbal and non-verbal inhibitory control. We further predicted: a) patients with left frontal tumours would exhibit significantly poorer scores (T scores) on our Stroop difference measures when compared to the remaining tumour groups; and b) patients with right frontal tumours would exhibit significantly poorer scores (T scores) on our each of our measures of nonverbal inhibitory control, when compared to the remaining tumour groups.

With respect to internally-driven generation, our predictions were as follows. Based on the hypothesis that this broader group of skills is also domain-specific (Robinson et al., 2012), we predict that there will be no association between our verbal and nonverbal measures of this skill. Further, we predicted that: 1) left frontal patients will perform significantly more poorly (in T scores) than right frontal patients on measures of verbal fluency; and 2) right frontal patients will perform significantly more poorly (in T scores) than left frontal patients on measures of design fluency.

Method

Participants

The sample of 28 neurological tumour participants was identical to that described in Chapter 2. The sample of healthy control participants was also identical to that described in Chapter 4.

Materials and Procedure

The tasks, materials, and procedure was identical to that described in Chapters 2 and 4. Appendix C provides a table of the task administration order during the long-term follow-up. Where possible, we attempted to alternate between computer tasks and tasks of other formats (e.g. oral and pen and paper) to minimise eye strain.

The operationalisation of each of our measures is described above, and summarised in Table 5.1.

Each patient's performance on each of our selected measures was converted to a z-score using the procedures outlined in Chapters 2 and 4. This standardisation method ensures that each of the measures were weighted equally, regardless of the measurement scale.

Table 5.1.

The domain/process and operationalisation for each of our verbal and non-verbal measures of inhibitory control and internally-driven generation respectively

Domain/ Process	Operationalisation
Verbal Inhibitory Control	a) Average response latency on incongruent trials of the Stoop task as a proportion of average response latencies on congruent trials, expressed as a z score relative to the entire patient sample
	b) The difference in the percentage of items correct on the incongruent and congruent conditions of the Stroop task respectively, expressed as a z score relative to the entire patient sample
Non-Verbal Inhibitory Control	c) Overall accuracy on the antisaccade task, expressed as a z score relative to the entire patient sample
	d) On the Choice and Prepare RT tasks, the average response times for letter A trials, expressed as a proportion of the average response times for the remaining three stimuli (B ,C, or D), expressed as a z score relative to the entire patient sample
	e) On the design fluency task (D-KEFS, Delis et al., 2001), the proportion of designs generated on the 'filter' condition, relative to 'basic' condition, expressed as a z score relative to the entire patient sample
Verbal Internally-Driven Generation	g) Total number of correct responses on the letter fluency task (FAS), expressed as a z score relative to the entire patient sample
	h) Total number of correct responses on the category fluency task (fruit, animals), expressed as a z score relative to the entire patient sample

	i)	On the BLAST verb generation task, the average number of incorrect items for nouns with *low response strength*, compared to the average number of incorrect items for nouns with *high response strength* (e.g., scissors -> *cut*), expressed as a z score relative to the entire patient sample
Non-Verbal Internally-Driven Generation	j)	Total number of correct designs on the "basic" condition of the design fluency task (D-KEFS, Delis et al., 2001), expressed as a z score relative to the entire patient sample
	k)	Total number of correct designs on the "filter" condition of the design fluency task (D-KEFS, Delis et al., 2001), expressed as a z score relative to the entire patient sample
	l)	Total number of correct designs on the "switch" condition of the design fluency task (D-KEFS, Delis et al., 2001), expressed as a z score relative to the entire patient sample

<center>**Results**</center>

Inhibitory Control

   Relationship between verbal and non-verbal material. Table 5.2 presents the summary data for each of the various measures of verbal and nonverbal inhibitory control. To detect significant impairment, each patient's z scores were compared to that of the relevant age-matched control sample using Crawford and Howell's (1998) modified t test. Blue cells represent significant impairment (p <0.05). As is evident from the Table, accuracy on the antisaccade task was the most sensitive at detecting significant impairment. However, the poor scores did not appear to be concentrated in any specific localisation group for any particular measure.

   As discussed previously, the verbal measures of inhibitory control yielded no significant impairment. Such findings are likely attributed to the poor performance of the healthy control participants on this task, which resulted in some floor effects. It should be borne in mind that the normative sample of the BLAST is relatively small (n = 20). Whilst Crawford and Howell's (1998) t test attempts to minimise effects of small control samples, a larger normative sample would nonetheless provide more representative data; under such conditions, greater significant impairments on verbal inhibition measures are likely to emerge.

<center>215</center>

Table 5.2.

Summary results of each patient's z scores on our selected measures of verbal and non-verbal inhibitory control. Averages and standard deviations (SD) are presented (in parentheses) for each tumour localisation.

	Antisaccade Accuracy	Filter	A vs Other	Stroop Latency Difference	Stroop Accuracy Difference
Left Frontal					
BP	0.792	-0.021	-0.150	-1.136	0.992
RF	0.792	1.371	-0.228	0.219	-0.169
TD	-0.076	1.098	-0.807	-0.978	-0.169
LC	0.358	0.824	-0.171	-0.867	0.992
DP	0.792	na	-0.686	-0.799	0.992
Average	*0.531 (0.388)*	*0.818 (0.602)*	*-0.408 (0.313)*	*-0.712 (0.536)*	*0.528 (0.636)*
Left Temporal					
CA	-0.728	0.277	-0.533	-0.291	0.992
SH	-1.379	0.642	-0.901	0.921	0.992
GM	-0.185	0.642	0.589	-0.573	-1.910
KW	0.792	0.025	-0.008	-0.467	0.992
Average	*-0.375 (0.918)*	*0.397 (0.301)*	*-0.213 (0.648)*	*-0.102 (0.692)*	*0.266 (1.451)*
Right Frontal				na	na
NH	0.683	-0.707	-0.768	0.771	0.992
DF	0.792	-0.191	-0.689	na	na
BS	0.792	-0.133	0.200	-0.160	-0.749
LW	-1.162	-0.379	0.343	-0.479	-0.169

GP	-1.270	-2.347	3.055	0.916	-0.749
RS	0.792	-0.816	-0.420	-0.189	0.411
SN	-1.053	-1.568	-1.346	0.714	-1.910
WR	0.683		-0.565	0.311	0.911
Average	*0.032 (0.991)*	*-0.877 (0.811)*	*-0.024 (1.355)*	*0.269 (0.551)*	*-0.181 (1.045)*
Left Parietal					
AVG	0.792	-0.191	0.787	-1.166	-0.749
BD		-0.953	-0.869	2.091	-1.330
CR	-1.162	0.277	-0.381	0.377	0.992
NOH	-0.945	0.277	0.366	-0.295	0.411
KG	0.358	-0.227	-0.280	-0.193	0.412
SO		-0.379	0.015	2.375	-0.169
Average	*-0.239 (0.960)*	*-0.199 (0.459)*	*-0.060 (0.585)*	*0.531 (1.410)*	*-0.072 (0.854)*
Right Posterior					
SG	-3.007	-2.066	0.951	-0.764	-0.749
AEK	-0.185	-0.816	0.441	-0.035	0.411
JB	0.792	0.642	0.369	-1.607	0.992
PC	0.792	0.824	0.964	-0.542	-0.749
Average	*-0.402 (1.796)*	*-0.354 (1.357)*	*0.681 (0.320)*	*-0.737 (0.655)*	*-0.024 (0.871)*
Undifferentiated					
AE	0.683	1.918	0.722	1.845	-1.910

Turning now to the main objective of this Chapter, which was to examine the relationship between verbal and non-verbal material. To do this, we performed a Pearson's correlation, using SPSS Software. Table 5.3 presents the uncorrected r values for the each of the selected measures.

As seen in the Table, there was a significant positive correlation between two non-verbal measures: accuracy performance on the antisaccade task and the proportion of designs on the filter condition. Such findings suggest that the oculomotor and behavioural inhibition required for the antisaccade task and filter condition respectively likely indexes a common underlying inhibitory process.

However, as predicted, there were no significant correlations between any of the verbal measures on the one hand, and the nonverbal measures on the other hand. These findings are consistent with the notion that prepotent response inhibition may indeed be subserved by two functionally distinct mechanisms – one that underpins verbal material, and one that underpins nonverbal material.

Table 5.3

Pearson's correlation coefficients for each of the verbal and non-verbal indices of inhibitory control (antisaccade acc = antisaccade accuracy).

		A vs Other	Antisaccade (Acc)	Filter - Simple	Stroop Latency Difference	Stroop Accuracy Difference
A vs Other	Correlation	-	-0.206	-0.228	-0.082	-0.302
	P Value		*.313*	*.263*	*.687*	*.125*
Antisaccade (Acc)	Correlation		-	0.485*	-0.161	0.171
	P Value			*.016*	*.433*	*.415*
Filter - Simple	Correlation			-	-0.147	0.130
	P Value				*.473*	*.536*
Stroop Latency Difference	Correlation				-	-0.329
	P Value					*.093*

*p<0.05

Tumour Localisation. To explore the effects of tumour localisation on each of our verbal and nonverbal measures, patients were categorised into one of the following five broad tumour groups: left frontal (n = 5), right frontal (n= 8), left temporal (n= 4), left parietal (n = 6), and right posterior (n = 6)[19] (see Chapter 2 for a description of these localisation classifications).

Our first prediction was that patients with left frontal lesions would exhibit significantly lower scores on our verbal measures of inhibitory control compared to the remaining tumour groups. To test this hypothesis, we first performed an omnibus one-way Analysis of Variance (ANOVA) using SPSS Software, considering all groups simultaneously, and then if there was a significant overall group effect, we specifically compared scores for the group of interest with those for the other groups considered together. However, inconsistent with this prediction, no significant effects of group were obtained, therefore we did not proceed with any more specific comparisons amongst subgroups.

Our second prediction was that patients with right frontal lesions would exhibit significantly lower scores on our non-verbal measures compared to the remaining tumour groups. A one-way ANOVA revealed a significant overall group effect for *Filter design fluency* ($F(4, 24) = 3.713$, $p= .020$). Consistent with our original prediction, a planned contrast revealed that those with right frontal lesions performed significantly more poorly than the remaining tumour groups when needing to "filter" out their previous responses ($t(20) = 2.870$, $p= .017$). No other effects of group were obtained.

As above, we also performed statistical power maps for each of the verbal and nonverbal measures to ascertain whether it was possible to proceed with voxel-based analysis. Again, due to the continuous nature of our data, we used the Wilcoxon-Mann-Whitney probability to specify whether there was sufficient statistical probability to detect damage to voxels in different brain regions. However, in all cases, there was insufficient power to detect a significant effect in any brain region at the familywise $p <.05$ false discovery rate. Therefore, in no instance did we proceed with more sophisticated voxel-based lesion analysis.

---

[19] One patient, AE, was excluded from this analysis due to the presence of multiple lesions

Internally-driven Generation

Relationship between verbal and non-verbal material. Turning now to our measures of internally-driven initiation. Table 5.4 presents the summary data and averages for each of the various measures and localisation groups. As is evident from the Table, there appears to be evidence of a dissociation in some patients' performance scores. Specifically, patient SH, who presented with a left temporal glioblastoma, showed differential patterns of performance on our verbal and nonverbal indices. Such observations are seemingly at odds with a domain-general account, however, to confirm this we would need to assess the relationship more statistically, which is discussed below.

Table 5.4.

Summary results of each patient's z scores on our selected measures of verbal and non-verbal internally-driven generation. Averages and standard deviations are presented (in parentheses) for each tumour localisation. N.B Response strength is taken from the verb generation task.

	Letter Fluency	Category Fluency	Response Strength	Basic	Filter	Switch
L Frontal						
BP	-0.636	0.347	0.023	0.798	0.519	0.895
RF	-0.502	-1.024	0.696	-0.698	-0.009	-0.358
TD	-0.770	-0.392	0.263	-0.100	0.519	0.268
LC	-0.033	0.557	-0.362	-0.698	-0.274	1.208
DP	1.440	1.507	-0.330	1.097	1.575	0.268
Average	*-0.100 (0.905)*	*0.199 (0.963)*	*0.058 (0.441)*	*0.080 (0.836)*	*0.466 (0.709)*	*0.456 (0.610)*
Left Temporal						
CA	0.368	-1.341	-0.074	-0.399	-0.274	0.268
SH	-1.507	-1.974	-0.518	0.200	0.519	-0.671
GM	-0.502	-0.814	2.843	0.200	0.519	-0.045
KW	-0.636	-0.392	0.343	1.397	1.047	0.581
Average	*-0.569 (0.768)*	*-1.130 (0.683)*	*0.648 (1.505)*	*0.349 (0.753*	*0.453 (0.545)*	*0.034 (0.535)*
Right Frontal						
NH	-0.301	1.085	-1.599	0.499	-0.274	0.268
DF	-0.368	-0.286	-0.518	-0.399	-0.538	-0.358
BS	0.033	1.296	-0.531	-0.100	-0.274	0.268

LW	-0.368	0.241	-1.019	-0.998	-1.066	-0.984
GP	-0.435	-0.708	1.103	-0.998	-1.858	-1.297
RS	0.971	0.557	0.343	1.097	-0.009	0.895
SN	0.100	-0.286	-2.140	2.295	-0.274	-0.984
WR	2.981	0.347	-0.330	-0.100	0.519	0.268
Average	*0.327 (1.168)*	*0.281 (0.695)*	*-0.586 (1.030)*	*0.162 (1.113)*	*-0.472 (0.716)*	*-0.240 (0.784)*
Left Parietal						
AVG	1.440	0.557	0.023	1.696	1.047	0.581
BD	-0.435	2.034	-0.346	-0.100	-0.802	-0.671
CR	2.043	0.557	0.023	0.798	0.783	-1.610
NOH	-1.172	-1.235	1.103	-0.698	-0.538	-0.671
KG	-0.435	-0.286	-1.409	1.397	0.783	2.147
SO	-0.100	-1.552	1.689	-0.998	-1.066	-1.923
Average	*0.223 (1.242)*	*0.013 (1.325)*	*0.180 (1.094)*	*0.349 (1.116)*	*0.035 (0.936)*	*-0.358 (1.508)*
Right Posterior						
SG	0.569	0.136	-0.411	-0.399	-1.594	0.895
AEK	-0.837	0.347	0.023	-0.698	-1.066	-0.671
JB	0.368	1.612	0.006	0.200	0.519	2.147
PC	-0.301	0.136	1.103	-0.698	-0.274	-0.671
Average	*-0.050 (0.643)*	*0.557 (0.710)*	*0.180 (0.647)*	*-0.399 (0.423)*	*-0.604 (0.925)*	*0.425 (1.365)*
Differentials						
AE	-0.971	-1.024	0.006	-1.297	-0.538	-0.358

The main objective of this Chapter was to examine the relationship between verbal and non-verbal internally-driven generation. To address this question, we performed a Pearson's correlation, using SPSS Software. Table 5.5. presents the uncorrected r values for this analysis.

As seen from Table 5.5., there were a number of significant positive correlations between our nonverbal measures of internally-driven generation. Specifically, for the total number of designs on the *Filter* condition, a significant positive correlation was observed between the number of designs produced on the *Basic* condition, and also on the *Switch* condition. Such findings are perhaps unsurprising and suggest that these measures are indexing a common, underlying process.

Further, as seen in Table 5.5., there was also a significant positive correlation between the letter fluency and category fluency conditions of the verbal fluency task. Again, such findings suggest that these conditions are likely indexing a common process.

Table 5.5

Pearson's correlation coefficients for each of the verbal and non-verbal indices of internally-driven initiation

		Letter Fluency	Category Fluency	Response Strength	Basic – Design Fluency	Filter – Design Fluency	Switch – Design Fluency
Letter Fluency	Correlation	-	0.429*	-0.128	0.326	0.328	0.125
	P Value		.023	.517	.091	.089	.525
Category Fluency	Correlation		-	-0.356	0.282	0.168	0.357
	P Value			.063	.147	.392	.062
Response Strength	Correlation			-	-0.386*	-0.055	-0.249
	P Value				.043	.781	.201
Basic – Design Fluency	Correlation				-	0.667*	0.352
	P Value					<.000	.066
Filter – Design Fluency	Correlation					-	0.425
	P Value						.024*

*p<.050

In terms of the relationship between verbal and non-verbal material, only one significant correlation was observed between our verbal measures on the one hand, and the nonverbal measures on the other hand. Specifically, a significant negative correlation was found between our verbal *response strength* measure (obtained from the BLAST Verb Generation task) and the number of designs on the *Basic* condition on the design fluency task. This finding was inconsistent with our original hypothesis of domain-specificity, and suggests that difficulty initiating designs on the *Basic* condition of the design fluency task was associated with better performance retrieving verbs that had low response strength. The relationship between these two measures is plotted in Figure 5.1. As is evident, this effect appeared to be driven by just one patient, S.N. who presented with a right frontal meningioma. Indeed, when this patient was removed from the correlational analysis, the effect was no longer significant ($R = -0.237$, $p= .234$).
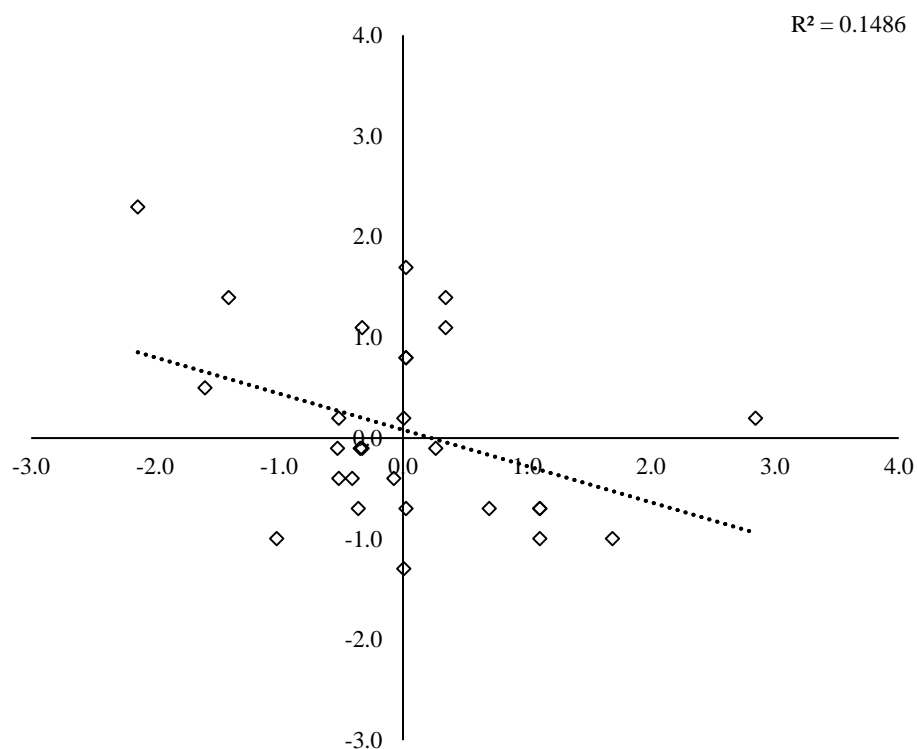
Figure 5.1. Scatterplot of the significant negative relationship between response strength and the total number of designs produced on the basic condition of the design fluency task.

Consistent with our original predictions, no other significant effects were obtained between our verbal and non-verbal measures of internally-driven generation. Such findings are broadly consistent with a domain-specific account of internally-driven generation.

Tumour localisation. To explore the effects of tumour localisation, patients were categorised into one of the following five broad tumour groups: left frontal (n = 5), right frontal (n= 8), left temporal (n= 4), left parietal (n = 6), and right posterior (n = 6) (see Chapter 2 for a description of these localisation classifications).

Our localisation predictions with regards to internally-driven generation are as follows. Patients with left frontal lesions would exhibit significantly lower scores on our verbal measures relative to the remaining tumour groups, whilst those with right frontal lesions would exhibit significantly lower scores on our nonverbal measures relative to the remaining tumour groups. As above, to test these predictions, we performed an omnibus one-way ANOVA, considering all groups simultaneously, and if a significant effect was obtained, we specifically compared scores for the group of interest with those for all other groups considered together. However, inconsistent with these predictions, no significant effects of group were obtained.

As above, we also performed statistical power maps for each of the verbal and nonverbal measures to ascertain whether it was possible to proceed with voxel-based analysis. However, in all cases, there was insufficient power to detect a significant effect in any brain region at the familywise $p <.05$ false discovery rate. Therefore, in no instance did we proceed with more sophisticated voxel-based lesion analysis.

Principal Components Analysis. To further explore underlying relationships between our various measures of internally-driven generation, we performed an exploratory Principal Component Analysis (PCA) using SPSS Software, and applying a varimax rotation. PCA is a method for examining the variance shared amongst different measures in order to extract the components that best explain the data. It also enables us to explore relationships amongst measures without the need to generate specific hypothesis.

For this exploratory analysis, we included each of our six measures of verbal and non-verbal internally-driven generation. These measures are already expressed in a z score format (as outlined above), which ensures that the variance associated with each measure is roughly equivalent. For this dataset, the Kaiser-Meyer-Olkin (KMO) measure – an estimate of the overall variance shared amongst variables - was 0.59, which although low, is considered an

adequate value of sampling adequacy (see Kaiser, 1974). Further, the Bartlett's test of Sphericity, which examines whether the observed correlation matrix significantly deviates from the identity matrix, also yielded a significant effect (χ^2 (15) = 40.24, p = <.000)[20]. Factors with an eigenvalue above 1 were extracted and then rotated using varimax rotation.

The PCA extracted two factors, which explained 61.1% of the total variance. Table 5.6 presents the factor loadings and the total variance explained by each individual factor. Values below 0.4 are not presented in the Table.

Table 5.6.

The factors extracted from our exploratory Principal Component Analysis (coefficients below 0.040 are not shown)

	Factor 1 (43.0%)	Factor 2 (18.2%)
Letter Fluency	.405	.427
Category Fluency	.	.813
Verb Generation Response Strength	.	.771
Basic Design Fluency	.805	.
Filter Design Fluency	.941	.
Switch Design Fluency	.554	.

As seen in Table 5.6, Factor 1 accounted for 43.02% of the total variance and contained strong, positive factor loadings for the total number of designs produced on the *basic* and *filter* condition of the design fluency task, and a moderate factor loading for the *switch* condition; this lower factor loading likely reflects the heavy involvement of a multitude of complex processes that are utilised during the switch condition. Overall, factor one predominately consisted of non-verbal conditions, and thus can be referred to as *non-verbal internally-driven generation*.

---

[20] None of our other investigations had an appropriate KMO rating or a significant Bartlett's test of Sphericity, therefore we did not proceed with exploratory PCA's in these cases.

Factor 2 accounted for 18.2% of the total variance and contained strong factor loadings for category fluency and our verb generation response strength measure. Accordingly, factor two was called *verbal internally-driven generation.*

Interestingly, letter fluency scores did not load strongly and specifically on either of these factors; scores on this measure were weakly loaded onto both factors.

In conclusion, whilst purely exploratory, the PCA is consistent with the view that there may be independent, domain-specific mechanisms serving the top-down, goal-driven control of verbal and nonverbal actions. However, it should be borne in mind that the number of participants in this sample was small (N=28). Some researchers have suggested that we should assess sample size relative to the number of variables entered into the analysis (sample to variable ratio; STV). Since there were 28 patients and six variables in our analysis, the sample-to-variable-ratio was 4:6. This value, although not uncommon in published studies, is considered to be at the low end of the acceptability range (see Fabrigar, Wegener, MacCallum, and Strahan, 1999; Costello & Osborne, 2003). Accordingly, the current findings from this exploratory analysis should be treated with caution and would need a larger sample size to be confirmed.

Discussion

The aim of this Chapter was to examine the functional organisation of two control processes that are proposed to play an important role in language functions, *inhibitory control,* and *internally-driven generation.* Specifically, we assessed whether these two classes of control processes operate in a domain-general or domain-specific manner – that is, are there functionally and anatomically distinct processes supporting control in the verbal and nonverbal domains. In the brief discussion below, we begin by considering those tasks we used to measure inhibitory control. We then turn to a consideration of those tasks which assess goal-driven action selection more broadly.

Inhibitory control. Our analysis of the selected inhibitory control measures revealed one interesting positive finding: there was a significant positive correlation between the proportion of designs on the "*filter*" condition of the design fluency task and accuracy on the *antisaccade* task. This finding suggests that the prepotent inhibition required for behavioural responses is the same as the oculomotor inhibition required on the antisaccade task. Such findings are in contrast to taxonomies that propose separate components of inhibition (see Nigg, 2000). Instead, this finding is broadly consistent with the notion of a general inhibitory control mechanism that operates over a range of nonverbal action domains. Further, this finding provides justification for combining non-verbal measures of inhibitory control into an overall aggregated measure (see Chapter 4).

No significant correlations were observed between our verbal measures of inhibitory control on the one hand, and non-verbal measures of inhibitory control on the other hand. This is broadly consistent with our original hypothesis, wherein we predicted that there are separate and distinct mechanisms of inhibitory control for verbal and nonverbal response domains. However, since this conclusion rests on a null finding, we should be careful in placing too much weight on it.

Finally, with respect to our anatomical hypotheses, only one of our predictions was supported by the data: there was a significant effect of localisation on scores for the filter condition of the design fluency task. Specifically, right frontal tumour patients scored significantly more poorly on this measure than the remaining tumour groups. Such findings support previous assertions that right frontal regions are involved in nonverbal aspects of inhibition.

The are several additional points to bear in mind when considering these findings. First, the absence of significant relationship amongst our verbal and nonverbal measures may reflect artefacts in the way our components were operationalised. For example, it could be argued that our measures of verbal and nonverbal inhibitory control index different types of inhibition – the suppression of *well-entrenched* responses in the Stroop task versus the suppression of *recently-learned* associations in the non-verbal tasks (Khng & Lee, 2009). Importantly though, whilst this may be a valid criticism for our Choice ('A' vs others) and Filter measures, this is possibly not the case for our antisaccade task, which requires the suppression of seemingly automatic and arguably entrenched saccades.

Second, as discussed in a previous Chapter, our measurement of the Stroop interference effect may have created artefacts not necessarily attributable to inhibitory processes. In the baseline ("congruent") stimuli on this task, the target ink colour was provided to participants within the word name itself (e.g., the word "red" written in red ink). Therefore, any difference in latencies between this condition and the incongruent condition may be as much due to a facilitatory effect of providing the target colour name in written form, as they are to interference between word and colour name. Therefore, direct comparison between incongruent and congruent trials may be confounded by facilitation effects, particularly for those who have difficulty selecting lexical items. Future studies could address this problem by including a neutral condition (for example, one where colour patches are presented and the colours must be named: see McLeod, 1991, for discussion).

Third, the absence of a significant correlation between the "filter" condition on the design fluency task and the remaining non-verbal measures may reflect different underlying processes. For example, it may be the case that the ability to "filter" out prepotent responses places demands not just on inhibitory control, but also on working memory. Of course however, a larger sample size would be needed to confirm these findings.

Other potential artefacts and limitations of the current study are discussed in more detail in the following chapter as they apply to the thesis as a whole. However, it is worth noting briefly here that our sample size was small, and we would of course need to examine a larger sample of tumour patients in order to draw firmer conclusions.

Internally-driven Generation. Turning now to our measures of internally-driven generation, significant positive correlations were found amongst our various measures of design fluency. Specifically, patients' scores on the *Filter* condition were significantly

associated with those on both the *Basic* and the *Switch* conditions. This finding suggests that all three tasks make use of a similar set of processes.

A significant positive correlation was also observed between our two verbal fluency measures –letter fluency and category fluency scores. Again, such findings suggest that the two types of demands likely index similar constructs of internally-driven generation.

We turn now to the main focus of this section – examining the relationship between verbal material on the one hand and non-verbal material on the other hand. Consistent with our original predictions, no significant correlations were observed between our measures. This was also the case when we performed an exploratory PCA. This PCA identified two major factors, one of which loaded on verbal measures, and the other on nonverbal measures. While the majority of measures preferentially loaded onto only these two factors, one exception was letter fluency, which loaded equally strongly onto both factors. It may be the case that letter fluency constitutes the most demanding of all these various fluency tasks (e.g. see Shao et al., 2014). Specifically, category fluency is similar to everyday production tasks (e.g. making a shopping list), and thus it has been argued that participants can better utilise existing associations between semantic categories to retrieve a response. Conversely, letter fluency does not make such demands, and instead relies heavily on retrieving items from a phonemic category – a feat that is seldom done in everyday speech production (Shao et al., 2014). Consequently, greater demands are placed on the retrieval system during this task. Further, letter fluency may also place demands on even higher aspects of cognitive control, such as performance monitoring, and working memory. Of course, our sample size was only marginally acceptable for conducting a PCA, so we would need to confirm these findings using a much larger participant sample.

One caveat to this finding is that all the fluency task measures used in this study – including both verbal and nonverbal - are timed. That is, any person who is slow at responding overall will score poorly on these tasks. Indeed, a supplementary, exploratory regression analysis revealed that speed of responding on button press tasks – as assessed using the activation measure from Chapter 4 – was a significant predictor of scores on both design and category fluency, but not scores on letter fluency. One way to tackle this potential confound in future is to factor out general speed effects statistically, by including a general response speed measure as a covariate in analyses. Another way to overcome this potential confound is to measure fluency not only by the number of correct responses, but by the timing of these responses. Specifically, Luo and colleagues (2010) have proposed that greater

demands are placed on cognitive control abilities across the duration of fluency trials as individuals need to monitor and remember their previous responses, suppress competing interference from these responses, and sustain their attention. In contrast, they suggest that the initial trials may be disproportionately more susceptible to difficulty with lexical access speed and retrieval mechanisms (Luo, Luk, & Bialystok, 2010). Accordingly, a more sensitive way to measure internally-driven generation may be to calculate a slope measure that describes the linear change in responses across the trial, based on the line of best fit.

Chapter 6:
General Discussion

The objective of this study was to examine the language and higher-level control capabilities in a sample of patients who have undergone tumour resection surgery. The major aims of the thesis were as follows:

1. To gain information about the long-term clinical impact of tumour and tumour resection surgery on language function.

2. To explore the relationship between performance on single word language tasks and sentence-level tasks in this population.

3. To assess the impact of tumour and tumour resection surgery on a range of tasks that require high-level cognitive control, and investigate the relationships between these skills.

4. To explore the relationship between the control of language function and other types of cognitive control processes that operate outside the verbal domain.

The main findings relating to each of these objectives are discussed below.

Study 1: Long-term language outcomes

The first aim was to investigate long-term language outcomes in an undifferentiated sample of 25 patients across three time periods: immediately pre-operatively (one day prior to surgery), immediately post-operatively (one to three days following surgery), and at least three months following surgery (mean = 5.7 months). In contrast to many previous studies, we used a cognitively-motivated assessment protocol, developed specifically for neurological tumour patients, called the BLAST. This assessment enables the researcher to calculate scores for eight "core" cognitive skills that are believed to be crucial for effective language comprehension and production. Using this approach, we found that a large proportion (49%) of our tumour patients had persistent language deficits at least three months after their surgery. The incidence of such difficulties was considerably higher than has been estimated in previous studies, particularly those that have used more conventional aphasia assessments validated primarily on post-stroke populations.

We also found that our core skills approach was generally very sensitive at detecting impairments in this population: the incidence of reported deficits was surprisingly high in this population across all surgical phases. Indeed, the core skills approach was considerably more sensitive at detecting impairment than were the raw scores (total correct) for each of the subtests we administered. Further, our core skills measures were also more effective predictors of performance at long-term follow-up than were the raw scores: at the subtest level, only *category fluency* was a significant predictor of follow-up performance, whilst at the core skills level, the following were significant - *accessing semantic knowledge, lexical selection, phonological encoding,* and *verbal short-term memory*. These findings suggest that subtle language deficits may be more common than previously estimated in tumour populations, and that these deficits may persist many months after tumour surgery. However, the deficits are likely to go undetected on more standard aphasia assessments. These findings provide significant new insights into the language capacities of tumour patients, and their associated long-term outcomes.

We also made a number of predictions as to the likely associations between tumour localisation and deficits on specific core skills. However, the approach we adopted here – which involved categorising patients into broad groups based on their primary lesion localisation - yielded very few significant differences for this small patient sample. Nevertheless, the core skill of *accessing semantic knowledg*e was approaching significance (p = .05), with left temporal patients scoring lower than the remaining groups. If confirmed in subsequent, larger studies, this finding would support the hypothesis that this region is involved in the processing of semantic information related to specific concepts and/or the mapping between this semantic information and their corresponding lexical labels. Unfortunately, we had insufficient power to assess the role of tumour localisation on a voxel-by-voxel basis.

Interestingly, lesion volume (cm^3) was *negatively* associated with the degree of long-term improvement on our language measures: patients with larger lesions tended to improve less than those with smaller lesions. The finding is contrary to our prediction, which was that those with larger tumour volumes would experience more relief from debulking surgery, and consequently, would enjoy greater improvement after surgery. Nevertheless, this finding is valuable in itself, because it may help us to predict which types of patients are most likely to require rehabilitation after surgery; this in itself can provide important cost-benefit ratios.

Finally, contrary to our predictions, chronological age and tumour grade were not significant predictors of long-term performance change, at least not in this small patient sample.

Study 2: Core skills and sentence-level speech

Our second aim was to investigate the relationship between connected, sentence-level speech comprehension and production, and the core cognitive language skills measured in the BLAST protocol. We reasoned that deficits involving the elemental cognitive skills assessed in the BLAST should impact upon more complex language performance in predictable ways. To assess connected speech production, we used the QPA (Saffran et al., 1989), which quantifies various properties of a spontaneously generated speech sample (e.g., rate of speech, proportion of open class vs. closed class words; proportion of words that advance the narrative). To assess connected speech comprehension, we used the TROG test (Bishop, 1989), in which patients must identify which picture best depicts an auditory sentence. The test requires participants to distinguish between different types of grammatical constructs and their associated meanings.

These types of sentence-level measures were not particularly sensitive at detecting significant impairment, at least not at long-term follow-up. This result may have been because at this stage, relatively few of our patients had significant deficits in the core skills that we hypothesised would be predictive of connected speech problems (e.g. auditory word recognition, verbal short-term memory). The finding may also reflect the lack of stringency in the connected speech assessments. For example, the QPA assesses only aspects of sentence form, not accuracy or the qualitative aspects of speech, and the TROG test provides cues to the listener that would not be available during real-time conversation.

Turning now to our predictions about the relationships between specific BLAST core skills and connected speech performance, few of our predictions were supported in this small sample. However, we did observe a significant association between BLAST *verb retrieval* scores and the QPA *struggle measure*, suggesting that the capacity to retrieve verbs from the lexicon is crucial for effective, fluent and concise speech. This observation may have important clinical implications as it suggests that interventions targeted at the individual verb level may also improve sentence-level language production. A significant correlation was also obtained between BLAST *lexical selection* scores and the two QPA measures of *closed-class words* and the *struggle measure*. These findings suggest that poor lexical selection skills manifest themselves in connected sentence-level speech in a number of ways: first, they result in speech that is associated with a greater proportion of false starts, repairs, and word repetitions. Second, lexical selection difficulties result in speech characterised by a higher proportion of closed-class function words, which is generally lacking in content words. Taken

together, these patterns will in turn affect the overall informativeness of the narrative. Again, this finding suggests that rehabilitation targeted at the individual lexical level may be beneficial in generalising to the sentence-level.

Study 3: Characterising Cognitive Control

Our third aim was to investigate the functional and anatomical organisation of higher-level cognitive control processes. To measure different aspects of higher level cognitive control, we again adopted a "core skills" approach. Drawing on contemporary theories of cognitive control, we defined and operationalised the following four components that we hypothesised are crucial for effective cognitive control: *activation*, *attentional maintenance*, *inhibitory control,* and *performance monitoring*. We then measured each of these core skills in a similar way to the BLAST: we identified multiple, theoretically-driven indices of each skill and aggregated each of these measures to create an overall skill score. This approach contrasts with that adopted in most previous studies, which have used highly complex and multifactorial tasks to assess cognitive control (see Chan, Shum, Toulopoulou, & Chen, 2008).

Using this approach, a considerable number of patients exhibited abnormally low scores on at least one of these core skills. Further, the absence of positive correlations amongst the various core skills measures was broadly supportive of our decompositional approach (that is, that higher-level cognitive control processes may be decomposed into several, more elemental processes). However, in our small sample, there were no statistically reliable associations between scores on each core skill measure and tumour localisation (defined on a coarse-grained, group basis). Again, a statistical power map revealed that there was insufficient power to perform any voxel wise analyses. However, it is worth noting that a number of individuals with posterior lesions also scored poorly on at least one of our measures. Since most previous studies examining "frontal" functions have been limited to frontal patients, relatively little is known about the involvement of posterior regions in the types of tasks often assessed in these studies. More research is needed that directly compares the performance of individuals with lesions to anterior and posterior cortical regions respectively. Nonetheless, the current study has made an important attempt at investigating the organisation of cognitive control, using very elemental measures.

Study 4: Domain-general versus domain specific language functions

Our final aim was to investigate the relationship between cognitive control of language processes and other types of control that operate outside the verbal domain. Specifically, we selected a range of tasks that required the inhibition of some well-learned action and/or goal-driven selection of a novel action. In some of these tasks, a verbal action was required (e.g., the Stroop task); in others, a nonverbal action was required (e.g. the antisaccade task). We assessed whether specific deficits and performance on the verbal tasks were associated with corresponding deficits and performance on their nonverbal analogues.

Overall, the absence of significant correlations between our verbal measures on the one hand and non-verbal measures on the other hand was broadly consistent with a domain-specific account of cognitive control - that there may be functionally distinct control systems operating on verbal and nonverbal material. There were few significant associations between individual task scores and (broad, group-based) lesion localisation. However, and consistent with our prediction, individuals with right frontal tumours scored significant more poorly than the other patients on a (nonverbal) design fluency task that involved overriding a previously learned design rule. Again, a statistical power map revealed that there was insufficient power to perform voxel-based analysis. Therefore, a larger sample is needed to confirm these findings.

General observations across studies

We did not directly examine the relationship between our various language core skills and the four cognitive control skills we identified in Chapter 4. First, with respect to associations between verbal processes and top-down cognitive control, it may be the case that our aggregated *activation* measure (described in Chapter 4) operates in a generalised manner, regardless of modality. For example, six patients exhibited significant impairments on this measure, with three of these (C.A., G.P., & S.O.) also showing significant impairment on the BLAST skill score *verb retrieval*. Curiously, each of these three patients had differential tumour pathologies and localisations. Instead, it may be the case that some patients' poor scores on *verb retrieval* reflect more general difficulties activating a response set to drive the retrieval process. Indeed, a number of theoretical frameworks emphasise that the recruitment of both linguistic and non-linguistic cognitive processes, such as energisation, is necessary to perform even basic language tasks, such as single-word tasks, as well as more complex operations involving sentence-level comprehension and production (Alexander, 2006; see also Bate et al., 2001; Shao et al., 2014). Our findings would appear to support this view.

Importantly though, as discussed in Chapter 4, our speed measures were relatively non-specific across the patient group as a whole (see Appendix F). In the current study, we have used very elemental measures of speed as a marker of a specific "activation/energisation process", outlined by Stuss and colleagues (1995; 2005). However, whilst there may indeed be such a process, it is worth considering that speed may be an outcome that reflects the efficiency of multiple processes rather than a specific PFC function. Of course, this has been difficult to determine in previous studies, which have selectively included *only* frontal patients. Therefore, the next step in this research would be to validate our measures of activation using a larger sample of undifferentiated tumour patients.

Exploration of the data also revealed some interesting observations involving our cognitive control core skill, *attentional maintenance*. When we consider individual patterns of performance, there appeared to be no association between this skill score and scores on our core language skills. For example, patient R.F., who presented with a left frontal meningioma, showed several significant impairments on our individual measures of activation, yet did not exhibit significant long-term impairment on any BLAST subtest or skill score, nor did he show deficits on any connected speech measure. Similar patterns of dissociation were evident across a number of patients who scored relatively poorly on our attentional maintenance measures, yet were largely unimpaired in the verbal domain or vice versa. Such observations may reflect the varying task demands of these constructs; specifically, our measure of attentional maintenance was based on sustaining attention during slow and infrequently visually presented stimuli. However, the verbal tasks did not place such demands as no stimuli were presented at a slow or infrequent rate. These findings appear consistent with previous assertions that attention is required for language processes to varying degrees, depending on the specific nature and demands of the task. Indeed, in a recent study of healthy participants, difficulties sustaining attention were most evident during more demanding, dual-task naming situations compared to a simple picture naming task (see Jongman, Roelofs, & Meyer, 2015).

Considering our four studies as a whole, it is apparent that a small number of patients showed generalised impairments across a large number of measures, regardless of modality. For example, patients S.O., G.M., S.H., A.E.K., & N.O.H., all of whom presented with high-grade tumours (e.g. glioblastoma and high-grade glioma), exhibited non-selective significant impairment across both the language and non-verbal core skills measures, irrespective of lesion location. Importantly, each of these patients had received at least one round of

chemotherapy and/or radiation prior to the follow-up assessment. Indeed, previous studies have reported a strong association between chemotherapy/radiation and long-term neurocognitive sequelea (see Douw et al., 2009; Taphoorn & Klein, 2004). For example, in an earlier cross-sectional study involving 104 patients with low-grade glioma, higher doses of radiotherapy were more likely to result in cognitive dysfunction than lower doses (Taphoorn, 2003). Further, radiation has long been associated with damage to the cerebral white matter, with side effects including subcortical dementia, tissue necrosis, demyelination and small vessel disease (Gehring et al., 2008; Monje, Mizumatsu, Fike, & Palmer, 2002; Sheline, Wara, & Smith, 1980; Welzel, Steinvorth, & Wenz, 2005). Whilst we attempted to mitigate treatment effects by assessing patients at least two months following their last treatment/ regime, the current observations suggest that there may continue to be long-term functional consequences, even some time after treatment has finished. Such information may be used to provide the patient greater capacity to make informed decisions regarding their treatment and likely outcomes. Importantly though, it remains unclear how these findings affect patients in their everyday lives. We have used very constrained tasks here; therefore, future studies should endeavour to assess the long-term functional impacts in an everyday context.

Absence of lesion specificity

As was evident throughout the four studies, very few lesion effects were evident. Due to the small size of our sample, it would be unwise to read too much into these null effects. However, in this context, it is perhaps worth noting that neurological tumours are space-occupying lesions that result in compression and/ or displacement of surrounding, non-affected neural tissue, and also have broad, generalised effects on cerebral metabolism and blood flow (Miceli et al., 2012). For this reason, tumours may impact on the functionality of tissue that is quite distant from tumour mass itself (Miceli et al., 2012). Indeed, evidence suggests that the mere occurrence of a tumour is sufficient to disrupt non-affected neural regions and pathways (McAleer & Brown, 2015). Further, the mass effects of a tumour may also give rise to the disruption of functions in the opposite, non-affected hemisphere (Giovagnoli, Casazza, Ciceri, & Avanzini, 2006).

Indeed, in a recent study, Bosma and colleagues (2008) reported differences in neural connectivity between 17 patients with LGG and matched healthy control participants (Bosma et al., 2008). Using short and long-range brain synchronization mapped to regions of working memory, attention, and information processing, the authors found a difference in long-range connectivity in the resting state of LGG patients; such differences led Bosma and colleagues

(2008; 2009) to conclude there was a possible mechanism related to overcompensation in patients with glioma, with indications that the focal presence of low-grade tumour in the brain affects overall functioning. Based on these observations, it is perhaps not surprising that there were very little anatomical effects in our sample. Indeed, rather than assessing lesion localisation, it may be more useful to examine tumour lateralisation (Gehring et al., 2008).

Advantages of the Current Study

An advantage of the current study is that it offers unique insights into a neurological tumour population more generally and extends our knowledge about long-term outcomes. As discussed previously, much of what is known about language processes comes from the study of stroke patients (Davie et al., 2009). However, the contrasting pathological mechanisms between tumour and stroke result in very different profiles (Anderson et al., 1990); therefore, it is necessary to further conceptualise language processes within tumour populations. Our study highlighted striking differences in the rate of impairment between our skill scores and previous studies that have assessed language using assessments largely validated on stroke patients. Moreover, the different rates of impairment at the core skills and sentence level may support previous assertions that tumour patients have unique language profiles, and thus should be assessed using assessments specifically tailored for this population. A further advantage of assessing neurological tumour patients is the unique distribution of lesion locations in this population. For example, in stroke aphasia populations, lesions tend to be concentrated around the distribution of the middle cerebral artery, with certain regions of prefrontal cortex, such as the anterior cingulate, relatively invulnerable to selective damage following a stroke. The non-selective localisations in the current study arguably allows for the examination of differential anatomical regions that are not typically impacted by vascular lesions. Accordingly, such evidence can complement the existing knowledge base that is derived primarily from stroke populations.

Another important feature of our study is the undifferentiated patient sample. Rather than selecting patients with tumours to specific regions known to be associated with language and/or cognitive control, we assessed a broad range of patients with lesions to anterior and posterior regions. This aspect of our methodology was unique from many previous studies of language and control functions, which selectively include left hemispheric and frontal patients respectively. Accordingly, the undifferentiated sample in the current study allows for greater insights regarding the validity of tasks commonly associated with left hemisphere and frontal regions.

The current investigation took a cognitively-motivated core skills approach to the examination of language and cognitive control. This approach has considerable advantages. Specifically, many previous studies of language and cognitive control abilities involve relatively blunt instruments: tasks that place multiple demands on the examinee and utilise highly complex material. For example, the Boston Diagnostic Aphasia Examination (Goodglass et al., 2001) is a widely-used assessment battery that explores a range of language capacities, such as free conversation, picture description, reading comprehension, sentence-level comprehension and narrative writing. Consequently, such an assessment places high demands on language processes that make it difficult to isolate the specific source of impairment. Our core skills approach offers a new way to explore the mechanisms that underpin cognitive and language functioning that is unique from many other studies. In this thesis, we also conducted a direct comparison between scores on our core cognitive language skills, and performance on sentence-level language tasks. This novel approach represents a new way of conceptualising and assessing cognitive function, not just in the language domain, but more generally.

Further, our approach of combining individual scores into aggregated, cognitively-motivated measures has several advantages over other commonly-used methodologies. First, by combining individual scores into their respective aggregated measure, we can better control for any noise variability relating to any specific measure. Second, this approach, which often involves making comparisons amongst different conditions or tasks – rather than simply examining overall accuracy or speed - limits the likelihood that performance would be confounded by extraneous variables, such as fatigue and concentration.

Of course, it is important to establish that the measures we use to index the various hypothesised core skills have validity. With respect to the core language skills assessed in the BLAST, a recent study has explored this issue, and the results appear to be very promising. Lee (2016) compared the core skills measured in the BLAST to scores on other neuropsychological tests that are hypothesised to index the same core skill. Overall, it was found that scores on the following skills were significantly associated with scores on an alternate, independent measure that was hypothesised to index the same theoretical construct: *accessing semantic knowledge, lexical selection, goal-driven response selection,* and *articulatory agility.* The next step in this research is to perform a similar validation study of the new cognitive control measures developed in this thesis. To do this, each of our core control skills would be compared against other, well-validated and commonly used measures

of *activation, attentional maintenance, inhibitory control,* and *performance monitoring.*

There were a number of additional strengths to our measures. First, each of the measures contributing to our core skill scores were first standardised relative to the entire patient group before being aggregated and then statistically compared with the relevant control sample (using Crawford & Howell's (1998) modified t-test). In the kinds of measures examined here, the distribution of scores in normal samples is often highly restricted, and may also be highly skewed. Therefore, if patient scores are standardised based on the means and standard deviations generated from this group, the values obtained can be very extreme. By standardising against the patient sample itself, we mitigated this problem, whilst still ensuring each of the measures contributing to each core skill were weighted roughly equivalently.

Second, when possible, individual performance scores were combined in such a way that they contributed to an aggregated overall "core" skill measure, which was then submitted for statistical analysis. This approach is useful for controlling for any confounding and extraneous variables that is more prone in individual raw scores, such as skewedness in the data and missing data samples.

Third, when comparing patients' core skill scores to healthy controls, we used Crawford and Howell's (1998) modified t-test, rather than using population-based statistics, such as means and standard deviations from our normative sample. This method is to be preferred when the control sample is small, as was the case here, because it treats the control sample as sample statistics rather than as parameters that describe the entire normal population. This method also has moderate power to detect a deficit, even when the normative population is small, and has also been shown to better control for type I and II error rates compared to other, similar tests (see Crawford & Howell, 1998; Crawford & Garthwaite, 2006)

Finally, each of our statistical analyses were based on an a-priori knowledge of theoretically-driven constructs and their associative predictions. This is important for ensuring that any interpretations of our findings is theoretically sound, whilst also limiting any premature interpretations of spurious results caused by our small sample size.

Limitations of the current study

Despite the advantages of the current study, there were a number of limitations that warrant discussion. The first concerns the small sample size. As stated throughout this thesis,

it is entirely possible that many of our non-significant findings reflect a Type II error due to a lack of statistical power rather than reflecting a genuine null effect. Related to this, it is worth noting that our significant correlations must also be treated with caution given they are likely to be powerfully influenced by one or two individual outliers, and thus may reflect a Type 1 error. A larger sample would allow us to ensure that these observed associations are secure. It would also enable us to perform additional, more sophisticated analyses, such as Principal Components Analysis, thus resulting in greater confidence in our data.

Also, although the number of participants in our studies appeared to be reasonable for a study of this kind (between 25 and 28 participants with neurological tumours), the sample was highly diverse. Lesion localisation was widely dispersed, and very few patients had overlapping lesions in any particular brain region. Consequently, we had insufficient power to perform any voxel-based lesion-behaviour analyses, and were restricted to comparisons across broadly defined lesion groups. This approach to lesion analysis is far from satisfactory, because few participants in each broad group are likely to share the same lesion localisation. Therefore, any genuine associations between lesion location and cognitive performance may be minimised and washed out by the grouping process. Whilst an undifferentiated sample is a strength of the current study as it enabled the identification of differential patterns of impairment, it is also a considerable limitation as it did not allow for substantial lesion overlap or subtraction methods that may have provided more specific localisation effects. A larger sample would enable us to perform finer-grained analyses of brain-behaviour relationships, such as voxel based lesion-symptom mapping.

A second limitation of the current study is the wide range of time intervals between surgical intervention and follow-up administration (3 - 12.2 months). Accordingly, there were likely to be considerable individual differences with respect to the opportunities for post-surgical recovery. Importantly however, this may not be as critical as it first appears. As discussed previously, in a sample of neurological tumour patients assessed over multiple follow-up sessions, Sanai and colleagues (2008) concluded that language recovery typically occurs by three months or not at all. Indeed, in their sample, there was very little change in language outcomes beyond three months.

A third limitation of the study is the absence of information regarding the location and extent of the surgical excision performed, and also information about tumour regrowth prior to long-term follow up. Indeed, whilst follow-up scans were able to rule this out in 44% of patients (11 of 25), this information was unavailable for the remaining sample. This is an

important limitation given regrowth may occur in both high and low-grade tumours. For example, although the survival rate for patients with meningioma is much higher than for those with high-grade tumours, the five-year mortality rate for persons aged between 55-65, is still 33% (Ostrom et al., 2017). These figures indicate that irrespective of tumour histology, regrowth is not uncommon. Accordingly, future studies should endeavor to obtain follow-up MRI scans of all patients to rule this out, which can ultimately limit confounding results.

Despite the advantages of the finer-grained measures that were used to infer cognitive control processes, there are several limitations with the current tasks that operationalise these measures. Many of the confounds and limitations that were present for our cognitive control measures have been discussed in previous chapters and are also detailed at length in Faulkner (2015; Faulkner et al., 2017). However, one important limitation worth mentioning concerns the operationalisation of BLAST core skill of *goal-driven response selection*. This skill score was derived from a combination of higher-level processes involving internally-driven action selection/ generation (*word-strength effect on verb generation* and *verbal fluency*) and pre-potent inhibition (*interference effect on the Stroop*). As such, it remains unclear how each of these processes relate to one another. Future studies should endeavour to use a larger sample size to examine the inter-relationships between each of the core skills measures to determine their coherence with one another. This can be achieved through more sophisticated statistical analyses, such as a principal components analysis (PCA).

A further limitation of the current study concerns the unique difficulties associated with neurological tumour patients when it comes to inferring brain-behaviour relationships. For example, Karnath and Steinbeck (2011) identified a number of problems associated with making such inferences in this population. First, they point out that we cannot assume that regions outside the tumour border remain intact. As discussed above, a number of studies suggest the spread of tumour boundaries can occur beyond those areas visualised in MRI (Burger et al., 1998; Kleihues et al., 2007; as cited in Karnath & Steinbeck, 2011). Karnath and Steinbeck (2011) also suggest that the brains of those with slow-growing tumours may be qualitatively different from those of healthy brains due to the process of functional reorganisation. Indeed, as discussed previously, a number of studies have revealed structural and functional differences between tumour patients and healthy controls, including intra-hemispheric connectivity (Bosma et al., 2008; Duffau, Denvil, & Capelle, 2002). Moreover, qualitative differences are also likely to emerge between tumour patients as a result of differential tumour pathologies and mechanisms. For example, those with low-grade tumour

are likely to have greater functional and structural reorganisation/ compensation given the slow-growing nature of their lesions. In contrast, patients with highly infiltrative, fast-growing lesions often do not exhibit the same degree of structural and functional reorganisation and recovery (Heimans & Reijneveld, 2012; Imperato Paleologos, & Vick, 1990; Klein & Heimans, 2004; Whittle et al., 1998). Accordingly, the decision to group our patients into an undifferentiated sample may have limited the detection of important differences that are driven by tumour grade and/or nature.

Of course, the above concerns are not limited to the study of neurological tumour patients, but apply to any neurological population (Shallice & Skrap, 2011). Indeed, the functional capacities and localisation of lesions within stroke populations is often identified using MRI techniques, which may also be inaccurate at detecting functional abnormalities in border regions. Moreover, the process of diachsis and functional reorganisation have also been observed in stroke patients (Green, 2003), thus there are also likely to be qualitative differences between the brains of stroke patients and healthy individuals. The question that remains to be answered is whether the magnitude of these concerns is greater and more serious in tumour, relative to stroke patients. Also, these disadvantages must be weighed carefully against the potential benefits of studying a neurological population that may offer unique information not obtainable from other sources.

Finally, the work in this thesis raises issues concerning the trade-off that we must always make between task transparency and ecological validity. In this thesis, we have opted for tasks and measures that are readily interpretable in terms of the component processes they involve. However, this transparency comes at a cost of ecological validity. They offer little in the way of predictions when it comes to a persons' ability to function within an everyday setting (for discussion, see Chan et al., 2008). Indeed, previous studies have found that individuals with lesions to frontal regions can perform comparably with healthy controls on traditional neuropsychological tests, yet experience significant difficulties in their everyday life (Goldstein, 1996; Shallice & Burgess, 1991). Although this thesis expressly addressed how the core, language skills assessed within the BLAST relate to performance on connected speech tasks, we did not extend our consideration to everyday conversation settings that involve more elaborate sentences and an understanding of the reciprocal nature of conversation. Similarly, in our investigation of core cognitive control skills, we did not address how scores on these measures impact on performance in everyday life. These questions do need to be addressed. However, we believe that the starting point of this

endeavour needs to be with the simple elemental measures of "core skills". Only once we have established appropriate operationalisation and validity of these concepts and measures can we then begin to assess their impact in everyday settings.

Conclusion

In conclusion, the current study examined language and cognitive control abilities in a sample of 28 undifferentiated neurological tumour patients at least three months after surgery. Our investigation adopted a cognitively-motivated approach, which aimed to identify and measure a series of "core" cognitive skills that are required for more complex behaviour in the domain of interest. This approach proved to have considerable sensitivity across the studies. Using this approach, we were able to show that language difficulties continue to be significantly impaired in neurological tumour patients well after the acute effects of tumour and surgery have diminished. Further, in the language domain, our exploratory investigations showed that there are likely to be consistent, reliable relationships between specific core skills and competency on sentence-level speech production and comprehension tasks. Our findings here provide unique insights into language processes and ultimately extend our understanding of the functional organisation of language, and the broader cognitive skills necessary for linguistic operations.

With respect to higher cognitive control, the core skills approach adopted here also highlighted dissociations between higher-level control skills, which supports decompositional accounts of cognitive control and suggests that the prefrontal cortex may be characterised by considerable functional and regional specialisation.

The patient sample studied here was relatively small, but the results obtained suggest that this type of approach to the assessment of cognitive function is highly promising, and worth exploring further in larger samples, and across different types of neurological populations. Such insights can ultimately expand our understanding of the mechanisms underpinning language functions, and the cognitive control processes that may be involved in, and necessary for, language functioning.

Appendix A:

Brain Tumour Patients Case Descriptions

This appendix provides a brief case description for each patient who participated in the long-term follow-up study. The Case descriptions have been categorised into each of the five broad anatomical groups. Each patient's brain scans, where available, are also presented on a standard MRI template (Rorden, et al. 2012) at MNI coordinates = -34, -24, -14, -4, 6, 16, 26, 36, 46, 56, 66. Some of these case descriptions are taken from Faulkner (2015), who completed pre and post-operative testing of 10 patients (see Chapter 2).

Left Frontal Group

BP

BP is an 18-year-old female from the North Island. In 2010, BP presented to her local hospital with a history of three complex grand mal seizures, which required hospitalisation. A subsequent MRI scan revealed a cystic lesion in the left frontal lobe that was consistent with a meningioma. MRI surveillance revealed an increase in the size and enhancement pattern of the lesion. A decision was then made to perform an elective craniotomy and resection of the tumour mass. BP underwent surgery in October 2014. There were no reported post-operative complications. Follow-up testing occurred four months later in February 2015. BP herself reported a full return to her previous activities, however she was continuing to take carbamazepine 800mg for seizure management. At the time of follow-up testing, a recent MRI scan had ruled out any tumour regrowth.
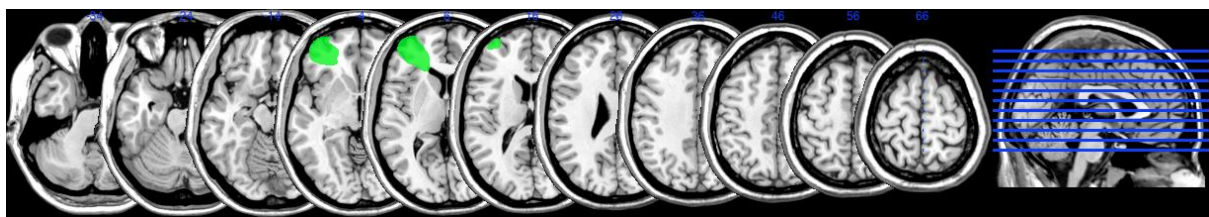
Figure A.1 MRI scan of patient BP

RF

RF is a 65-year-old male from the South Island of New Zealand. RF was admitted to Wellington Hospital in July 2012 after experiencing a partial seizure that resulted in a loss of consciousness. A subsequent MRI scan revealed that RF had a meningioma in the left parafalcine, posterior frontal lobe. RF's surgery was performed in July 2013. RF reported that he did not experience any post-surgical, nor long-term complications as a result of his

surgery. RF completed long-term follow-up testing 12 months later in July 2013. At the time of follow-up testing, RF was retired from his labouring job and a recent MRI scan had ruled out any tumour regrowth.
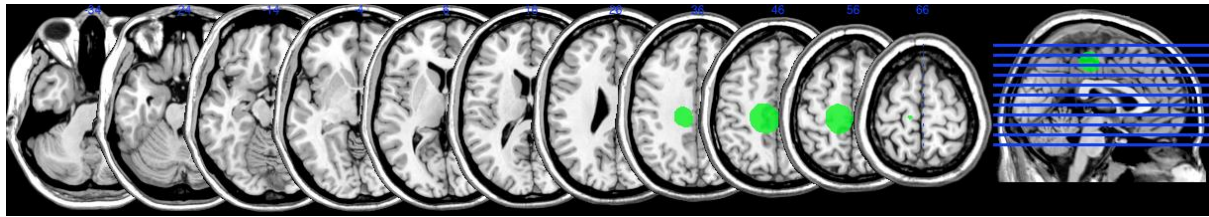
Figure A.2. MRI scan of patient RF

TD

TD is a 30-year-old male from the North Island. He experienced a three-month history of right-sided focal seizures. An MRI scan revealed a large generally non-enhancing mass in the left frontal region extending across the corpus callosum. The appearance was suggestive a low-grade glioma. Craniotomy occurred for debulking of the tumour in September 2013. Follow-up assessment occurred eight months later in May 2014. At the time of follow-up assessment, TD had re-commenced his employment and did not report any long-term effects of surgery.
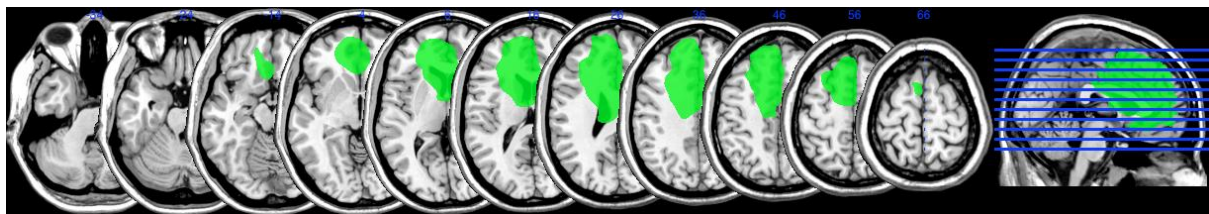
Figure A.3. MRI scan of patient TD

LC

LC is a 43-year-old female from the North Island. She experienced radicular pain, which had been present over the last few months. On admission into the Neurosurgical Ward in September 2013, she remained asymptomatic except for occasional headaches. LC's neurological examination was normal, with the exception of radicular pain in her right leg and reduced sensation on her right foot. An MRI scan on admission revealed a low- grade astrocytoma in the left posterior frontal lobe. In September 2013, an awake craniotomy was performed. Follow-up assessment occurred three months later in December 2013. At the time

of follow-up assessment, LC had not re-commenced her employment in healthcare due to the effects of fatigue.
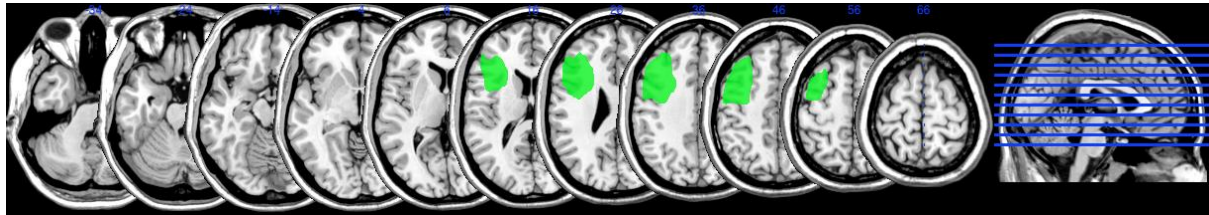
Figure A.4. MRI scan of patient LC

DP

DP is a 30-year-old male from the North Island of New Zealand. In June 2015, DP presented to Wellington Hospital with a two-month history of seizures. A subsequent MRI scan revealed a left frontal mass that was consistent with a low-grade astrocytoma. A functional MRI confirmed that the tumour was impinging on verb generation areas. DP underwent an elective awake craniotomy in June 2015. He experienced no post-surgical complications. Follow-up testing occurred 12 months later in June 2016. At the time of follow-up assessment, DP reported no long-term complications as a result of his surgery and had resumed full-time work.
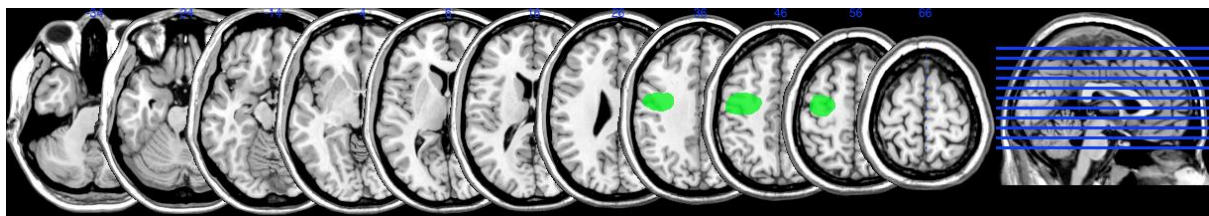
Figure A.5. MRI scan of patient DP

Left Temporal Group

CA

CA is a 47-year-old female from the North Island. In February 2013, she experienced a severe headache and deteriorating cognitive function. Subsequent neuroimaging revealed a large left sphenoid wing meningioma causing considerable midline shift and extensive central oedema. CA was commenced on Dexamethasone and underwent craniotomy and tumour resection in early March 2013. Follow-up testing occurred four months later in July 2013. At the time of follow-up assessment, CA had resumed her work as a nurse on reduced hours and continued to complain of headache, fatigue, and concentration difficulties.
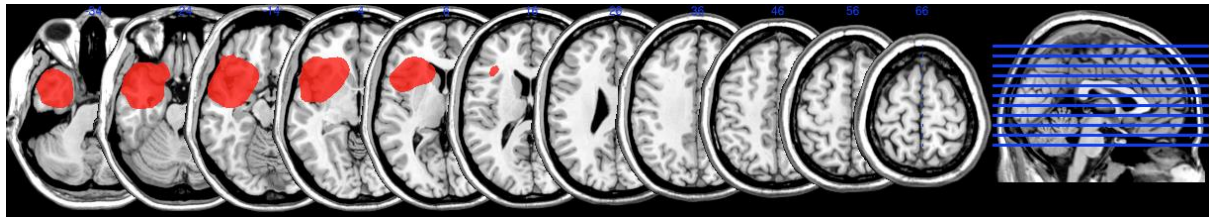
Figure A.6. MRI scan of patient CA

GM

GM is a 70-year-old male from the North Island. In early July 2013, GM experienced a two-week history of headaches and visual field deficits following two minor head injuries at work. An MRI scan revealed an irregular enhancing mass in the left temporal/ occipital pole that was consistent with a glioblastoma. Surgery was performed one day later to remove the bulk of the tumour. GM received a course of radiation therapy and was seen three months after his surgery in September 2013 for a follow-up assessment. At the time of follow-up testing, he continued to experience post-surgical fatigue, and subsequently had not resumed his work as a heavy manual labourer.
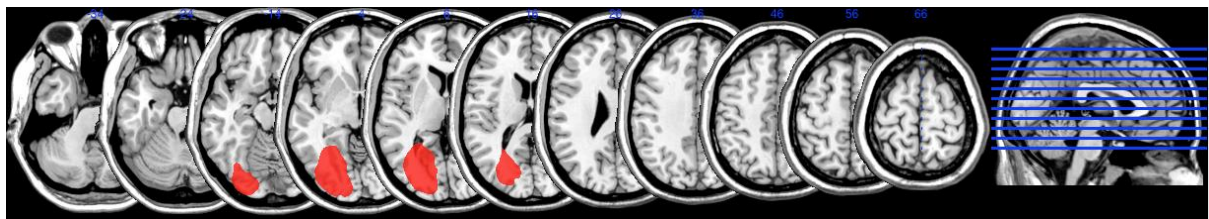
Figure A.7. MRI scan of patient GM

SH

SH is a 55-year-old male from the North Island. In September 2014, SH presented to hospital with a three-week history of headaches, visual blurring, and transient speech disturbance and smell. An MRI scan revealed a left temporal mass, consistent with a glioblastoma. SH underwent surgery one day later to remove the bulk of the tumour. He experienced post-operative seizures and expressive dysphasia. SH commenced chemotherapy and was seen in January 2015 for a follow-up assessment – two months after his last course of treatment. At the time of follow-up, SH continued to experience expressive aphasia, particularly in the context of word finding difficulties. He had not resumed his office job due to the effects of fatigue.
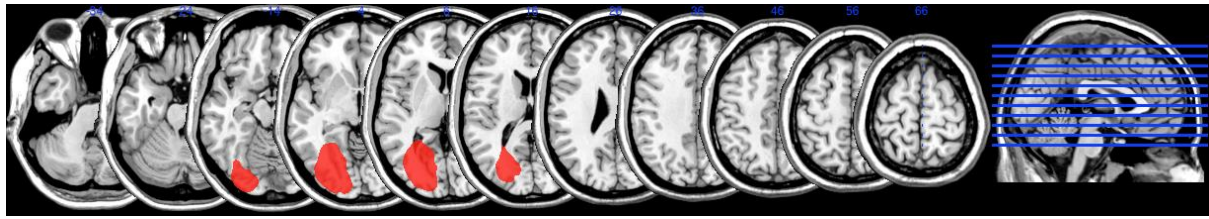
Figure A.8. MRI scan of patient SH

KW

KW is a 21-year-old male from the North Island. KW had a background of a left temporal cystic/ solid dysplastic lesion that was consistent with a meningioma (diagnosed in 2010). Routine MRI scans revealed that KW's lesion was increasing in size, with cystic changes. KW was subsequently admitted to Wellington Hospital in October 2014 due to a history of increasing headaches and three grand mal seizures that year which had required hospitalisation. In late October 2014, KW underwent an elective craniotomy and debulking to remove the tumour mass. KW experienced post-operative partial seizures, and was commenced on dexamethasone and sodium valproate (500mg). KW's follow-up assessment occurred four months later in February 2015. He reported continued seizures, however had resumed part-time working hours.
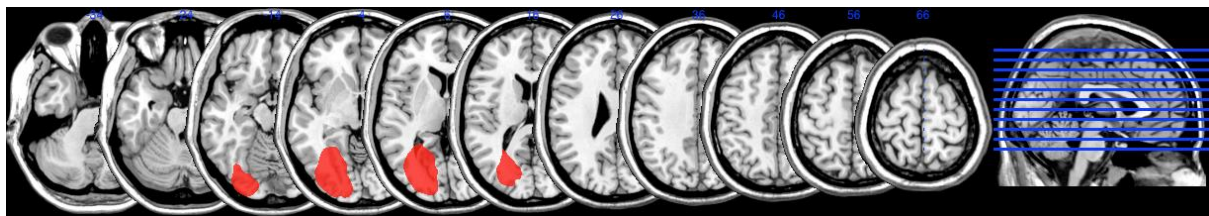
Figure A.9. MRI scan of patient KW

Left Parietal Group

AVG

AVG is a 42-year-old female from the North Island. In late March 2013, AVG experienced a sudden decline in her right hand function, symptoms of dysphasia, and word-finding difficulties. An MRI scan revealed a large durally based mass in the left convexity that was suggestive of a meningioma. Surgery was performed ten days later to remove the bulk of the tumour. Follow-up testing occurred five months later in August 2013. AVG reported no post-surgical or long-term complications and had resumed work full-time as an

administrator. At the time of follow-up testing, a recent MRI scan had ruled out any tumour regrowth.
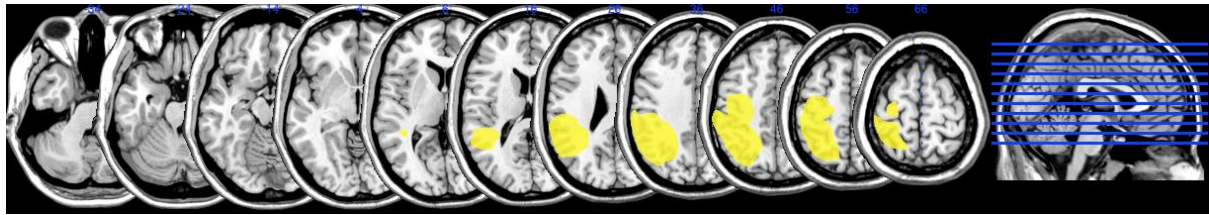
Figure A.10. MRI scan of patient AVG

BD

BD is a 62-year-old female from the North Island. She experienced symptoms of discomfort and reduced sensation in the proximal right lower limb, which had been present for the past year. BD presented to the Neurosurgery Ward in March 2013 and an MRI scan revealed a 3 x 4cm meningioma arising from the superior sagittal sinus. Craniotomy and resection occurred in March 2013. Follow-up testing occurred five months later in August 2013. At the time of this testing, BD continued to experience post-surgical fatigue, and subsequently had not resumed work as an administrator.
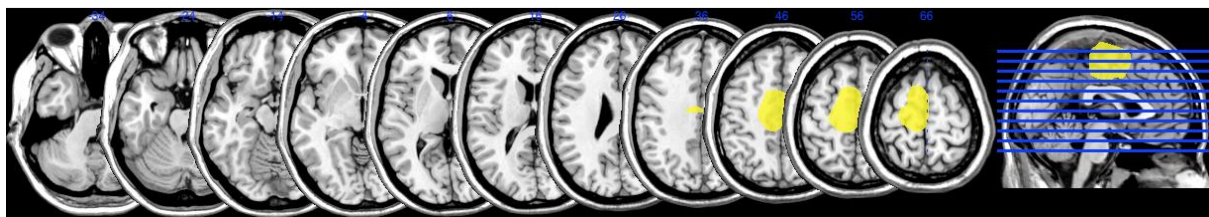
Figure A.11. MRI scan of patient BD

SO

SO is a 58-year-old male from the North Island. SO experienced a three-week history of declining right upper limb function. An MRI scan in August 2012 revealed an irregular heterogeneously enhancing tumour just right to the anterior motor strip in the left, posterior hemisphere. This was presumed to be a glioblastoma. SO underwent surgical debulking in late August 2012, followed by one round of chemotherapy and radiation. Follow-up testing occurred seven months later in March 2013, two months after the completion of his chemotherapy and radiation. SO required two separate testing sessions due to fatigue. At the time of follow-up testing, a recent MRI scan had ruled out any tumour regrowth.
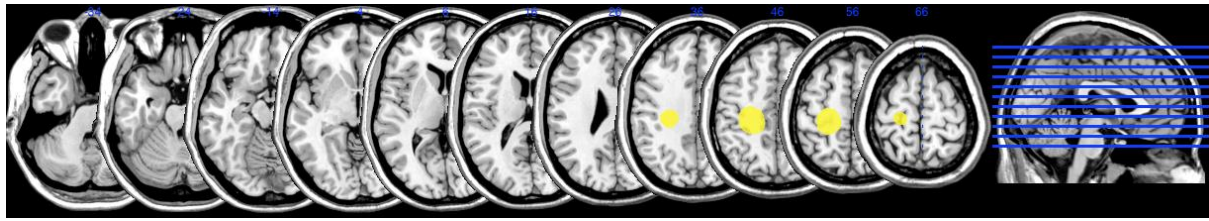
Figure A.12. MRI scan of patient SO

CR

CR is a 75-year-old female from the North Island. She experienced a period of weakness in her right hand, tightness under the chin, and dribbling on her right side. There was no disturbance of speech production, comprehension or expression. An MRI scan in April 2013 showed a 4.5cm left posterior tumour with homogenous enhancement suggestive of a meningioma. Craniotomy occurred in May 2013. Follow-up testing occurred four months later in September 2013. At the time of the follow-up assessment, CR reported no postsurgical or long-term complications as a result of her surgery.
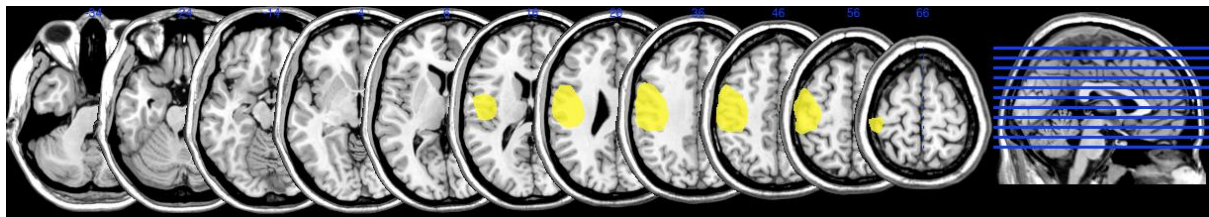
Figure A.13. MRI scan of patient CR

NOH

NOH is a 76-year-old female from the North Island. She presented to Wellington Hospital in early June 2014 after experiencing a seizure with transient dysphasia. NOH also experienced intermittent word-finding difficulties. A subsequent MRI scan revealed an enhancing peripheral tumour in the left posterior lobe with significant oedema. NOH underwent an elective craniotomy and resection for a high-grade glioma. During the acute post-operative stage, she continued to experience word-finding difficulties and was noted to have some mild expressive aphasia. She was commenced on sodium valproate. Follow-up testing occurred three months later in September 2014. NOH reported that her word-finding had substantially improved, although she continued to experience fatigue and was still taking sodium valproate for seizures.
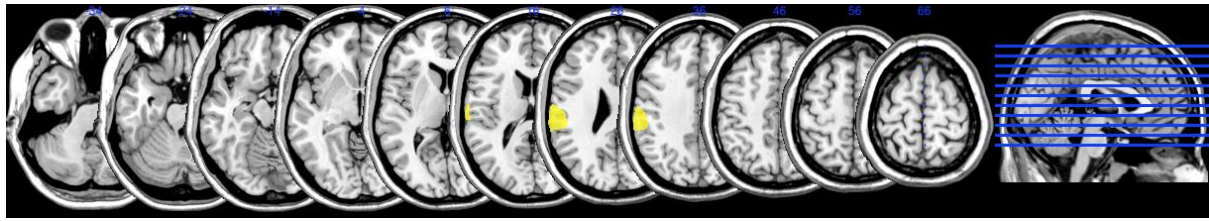
Figure A.14. MRI scan of patient CR

KG

KG is a 41-year-old female from the North Island. In 2014, KG reported several months of focal seizures affecting her right arm, including progressive weakness in her right elbow and some impaired coordination in her right hand. A subsequent CT and MRI scan revealed a non-enhancing lesion just anterior to the motor strip in the left hemisphere which was consistent with a low-grade glioma. FET and PET scanning confirmed motor activation in the gyrus just posterior to the tumour. In May 2014, KG underwent an elective craniotomy and debulking under awake conditions to enable the position of the tumour to be mapped. There were no reported post-operative complications. Follow-up testing occurred four months later in September 2014. KG reported continued difficulties with her right hand, however she had been able to resume work as an office administrator. At the time of the follow-up assessment, KG was taking sodium valproate for seizure control. A recent MRI scan had ruled out any tumour regrowth.
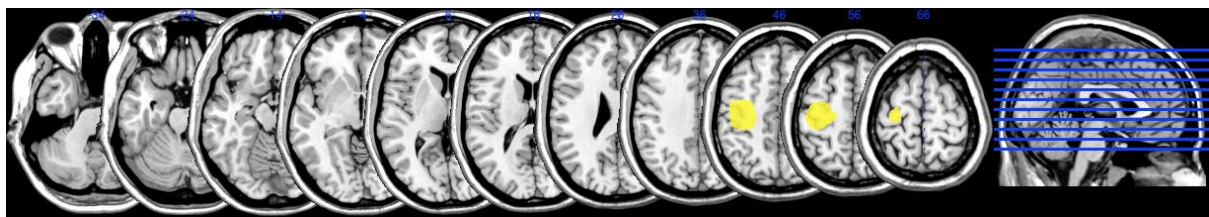
Figure A.15. MRI scan of patient KG

Right Frontal Group

BS

BS is a 56-year-old female from the North Island. In September 2014, BS experienced a seizure resulting in a loss of consciousness. She also complained of intermittent left hand paraesthesia and some numbness to the right side of her face. There were no reported

weaknesses or ataxia, and no visual disturbances, speech problems, or cognitive and personality changes. A subsequent MRI scan revealed a right hemispheric mass on the medial sphenoid ridge that was consistent with a meningioma. In October 2014, BS underwent an elective craniotomy for resection of the mass. There were no reported post-operative complications. Follow-up testing occurred four months later in February 2015. BS reported no long-term complications, and had resumed full-time work in an office.
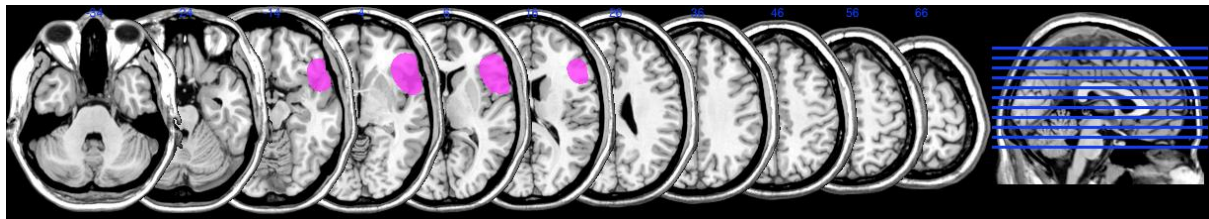
Figure A.16. MRI scan of patient BS

NH

NH is a 42-year-old female from the North Island. NH presented to Wellington Hospital after an 18-month history of visual deterioration and headache. A subsequent MRI scan revealed a right anterior clinoidal region tumour consistent with a meningioma. In September 2014, NH underwent an elective craniotomy for resection of the tumour mass. There were no reported post-surgical complications. Follow-up testing occurred three months later in December 2014. At the time of follow-up testing, NH reported no long-term complications and had resumed work full-time as an office administrator.
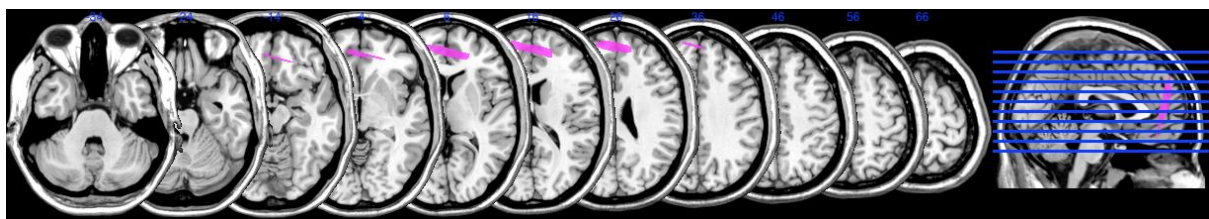
Figure A.17. MRI scan of patient NH

LW

LW is a 66-year-old female from the North Island. In October 2011, LW started to experience severe right-sided headaches. A subsequent MRI scan in June 2012 revealed an extra axial durally based lesion in the right posterior frontal convexity region, which was

consistent with a meningioma. It had a well-defined border but a somewhat lobulated appearance. There was a small amount of associated oedema. LW underwent a surgical excision of the tumour in August 2012 followed by chemotherapy due to primary lung cancer. Follow-up testing occurred in July 2013, four months after the completion of LW's therapy and 11 months after her surgery. LW required two separate testing sessions due to fatigue. At the time of follow-up testing, a recent MRI scan had ruled out any tumour regrowth.
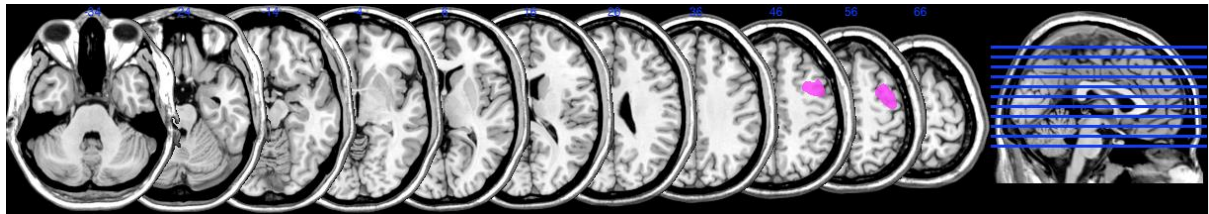
Figure A.18. MRI Scan of patient LW

DF

DF is a 40-year-old male from the North Island. DF experienced a history of several focal seizures, but no associated loss of consciousness. A subsequent MRI scan in May 2013 revealed a low-grade astrocytoma in the opercular, frontal and temporal region extending into the insular cortex, just shy of the external capsule. Craniotomy occurred for debulking in November 2013. Follow-up testing occurred eight months later in July 2014. At the time of the follow-up assessment, DF reported no long-term complications and had resumed full-time work as a tradesman.
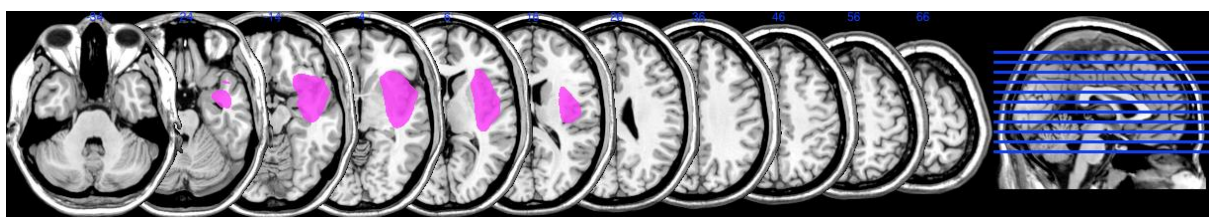
Figure A.19. MRI scan of patient DF

GP

GP is a 73-year-old male from the North Island. In November 2013, he presented to the Neurosurgical ward of Wellington Hospital with a several-month history of lethargy, disturbed gait and impaired coordination. A subsequent MRI Scan revealed a large right frontal tumour with considerable oedema and associated midline shift. Appearances were consistent with a large meningioma. Craniotomy occurred one week later for resection of the

meningioma. Follow-up occurred three months later in February 2014. GP continued to experience fatigue, but his gait and coordination had returned to normal following the surgery.
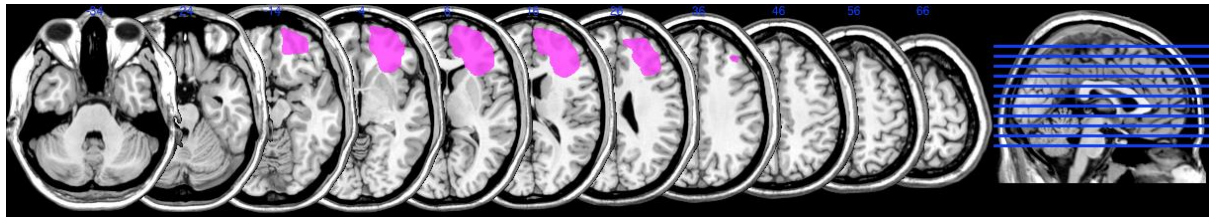
Figure A.20. MRI scan of patient GP

SN

SN is a 74-year-old male from the North Island. In June 2014, SN presented with a six-month history of a droopy eyelid and right sided temporal pain. A subsequent MRI scan revealed an en plaque tumour extending across the sphenoid wing into the temporalis muscle and anterior middle fossa. A transorbital biopsy confirmed a meningioma. In June 2014, SN underwent craniotomy and resection. Follow-up testing occurred four months later in October 2014. SN reported no post-operative or long-term complications as a result of his tumour/surgery and was able to resume his activities of daily living. There was no preoperative MRI scan available for SN, however, at the time of follow-up testing, a recent scan had ruled out any tumour regrowth.

WR

WR is a 60-year-old male from the North Island. WR had a background of a frontal olfactory meningioma, which had been under regular surveillance since 2010. In 2015, WR presented to an outpatient clinic with a six – eight month history of increased light-headedness and dizziness. A subsequent MRI scan in January 2015 revealed the mass had increased in size compared to the previous scan in September 2010. WR underwent an endoscopic resection of the tumour in March 2015. Follow-up testing occurred 12 months later in March 2016. WR reported no long-term complications and had resumed work in a labouring role. At the time of follow-up testing, a recent MRI scan had ruled out any tumour regrowth.
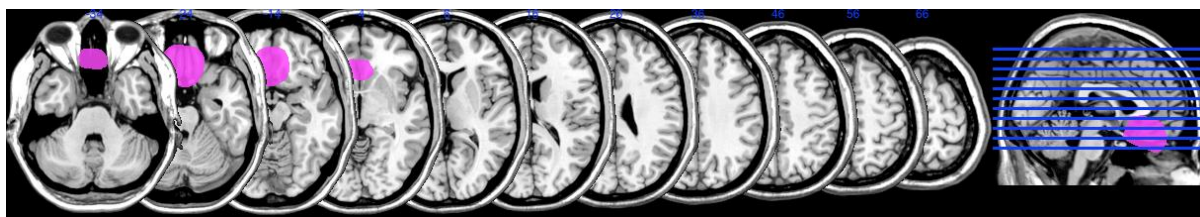
Figure A.21. MRI scan of patient WR

RS

RS is a 66-year-old male from the North Island. RS was admitted to Wellington Hospital in late September, 2014 following a seizure at home. A subsequent CT and MRI scan revealed a right middle cranial fossa enhancing mass that was consistent with a meningioma. In early October 2014, RS underwent an elective craniotomy and resection. There were no reported post-operative complications, with a subsequent MRI scan showing no evidence of a residual tumour. Follow-up testing took place eight months later in June 2015 at RS's place of work. He reported no long-term complications of the surgery and had commenced full time employment as an executive.
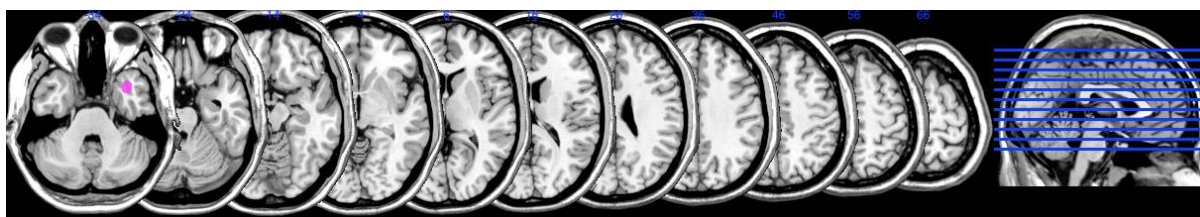
Figure A.22. MRI scan of patient RS

Right Posterior Group

AEK

AEK is a 52-year-old female from the North Island. She presented at Wellington Hospital with left facial drop and complaints of lethargy for the previous few months. A CT scan and subsequent MRI revealed a right posterior temporal tumour suggestive of malignant glioma. AEK underwent craniotomy and resection of the right fusiform gyrus malignant glioma in March 2013. AEK also underwent a course of chemotherapy and radiation and follow-up testing took place four months after her last course in July 2014.
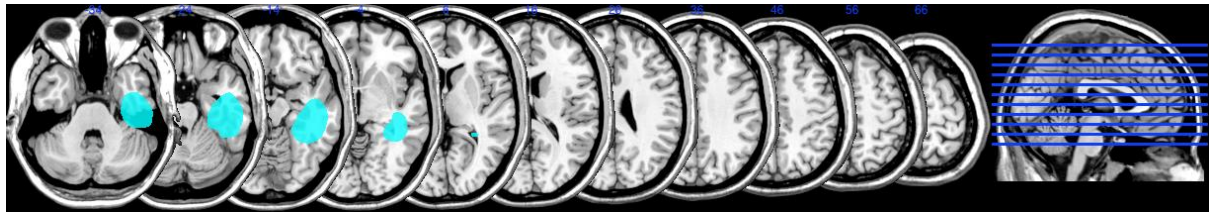
Figure A.23. MRI scan of patient AEK

JB

JB is a 47-year-old female from the North Island. Limited medical records were available, however she presented with a right temporal lesion that was consistent with a meningioma. JB's surgery was conducted in May 2012 and follow-up testing occurred 11 months later in April 2013. JB reported that she did not experience any post-surgical or long-term complications. At the time of follow-up testing, JB had resumed full-time work in healthcare. A recent MRI scan had ruled out any tumour regrowth.
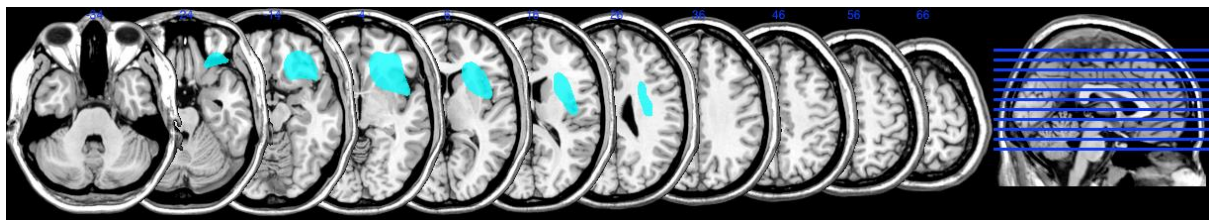
Figure *A.24.* MRI scan of patient JB

PC

PC is a 34-year-old male from the North Island. He presented to Wellington Hospital in June 2014 with a year-long history of worsening headaches and seizures affecting his left side. A subsequent MRI scan revealed a right parietal-occipital convexity lesion that was suggestive of a meningioma. PC underwent a craniotomy and resection to remove the bulk of the tumour. There were no reported post-operative complications. Follow-up testing occurred three months later in September 2015 at PC's place of work. PC reported no long-term complications and was able to resume full-time work in an office setting. Unfortunately, there was no MRI Scan available for PC.

SG

SG is a 42-year-old male from the North Island. SG had a background of a right occipital low-grade astrocytoma, which was diagnosed in 2010. A routine MRI scan in mid-2014 showed thickening and increased enhancement, which was indicative of tumour

progression. SG experienced intermittent vacant seizures lasting for approximately 10 minutes. He was admitted for a craniotomy and resection in July 2014 and experienced no post-operative complications. Follow-up testing occurred in November 2014. SG reported that he continues to experience intermittent seizures that affect his ability to resume work as a labourer. There were no other complications as a result of surgery. Unfortunately, there was no preoperative MRI Scan available, however at the time of follow-up testing, a recent MRI scan had ruled out any tumour regrowth.

Multiple Lesions

AE

AE is a 46-year-old man from the North Island. He presented to Wellington Hospital after experiencing complex partial seizures. Subsequent neuroimaging showed multiple intracranial cavernous meningioma and a right posterior frontal meningioma. The left posterofrontal lesion had increased in size, and was therefore removed by craniotomy in April 2013 whilst the right posterior frontal meningioma remained in situ at the time of follow-up testing. Follow-up testing occurred four months later in August 2013. AE reported no post-surgical or long-term complications although was no longer able to participate in sporting activities or driving due to his seizure history.
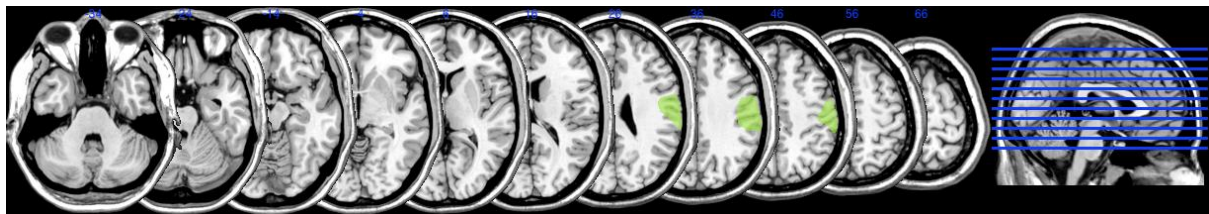
Figure A.25. MRI scan for patient AE

Appendix B:

BLAST Method and Materials

Table B.1

Frequency and Length of Items in Version 1 of the picture naming task in order of appearance

Item	CELEX Lemma Frequency	Frequency Category	Syllable Length
watch	710	hi	mono
camel	449	hi	mono
hamburger	86	med	poly
dinosaur	93	med	poly
apple	546	hi	mono
strawberry	115	med	poly
Kilt	34	lo	mono
Saw	62	lo	mono
monkey	324	hi	bi
carrot	144	med	bi
balloon	112	med	bi
parachute	7	lo	poly
ladder	287	hi	bi
hospital	2300	hi	poly
clown	65	lo	mono
Nest	304	hi	mono
cucumber	85	med	poly
butter	490	hi	bi
wheelbarrow	22	lo	poly
pyramid	123	med	poly
binoculars	93	med	poly
hoof	137	med	mono
guitar	119	med	bi
shark	357	hi	mono
scarecrow	18	lo	bi
cigarette	1274	hi	poly
astronaut	50	lo	poly
crutch	73	lo	mono
vegetables	1050	hi	poly
tongs	29	lo	mono
Crab	170	med	mono
lipstick	129	med	bi
chicken	734	hi	bi
apron	164	med	bi
caterpillar	58	lo	poly
Owl	128	med	mono
coconut	51	lo	poly
finger	2212	hi	bi
genie	16	lo	bi
telescope	142	med	poly
whale	199	med	mono
tomatoes	255	hi	poly

Item	CELEX Lemma Frequency	Frequency Category	Syllable Length
cake	610	hi	mono
drill	141	med	mono
sandwich	247	hi	bi
reins	13	lo	mono
hippopotamus	24	lo	poly
envelope	439	hi	poly
Raft	69	lo	mono
banjo	8	lo	bi
well	165	med	mono
submarine	311	hi	poly
lighthouse	50	lo	bi
necklace	71	lo	bi
chair	2441	hi	mono
rhinoceros	30	lo	poly
cannon	109	med	bi
skirt	522	hi	mono
goat	506	hi	mono
igloo	14	lo	bi

Table B.2

Frequency and length of items in version 2 of the picture-naming task in order of appearance

Item	CELEX Lemma Frequency	Frequency Category	Syllable Length
flower	1674	hi	bi
pocket	1343	hi	bi
pillow	344	hi	bi
stilts	18	lo	mono
potato	639	hi	poly
Waterfall*	137	med	poly
refrigerator	187	med	poly
typewritter	2300	hi	poly
elephant	429	hi	poly
pear	112	med	mono
penguin	90	med	bi
ambulance	162	med	poly
library	1113	hi	poly
pipe	558	hi	mono
hammer	197	med	bi
scissors	79	med	bi
pencil	332	hi	bi
volcano	102	med	poly
wreath	63	lo	mono
bottle	2079	hi	bi
cauliflower	43	lo	poly
pumpkin	38	lo	bi
jockey	95	med	bi
veil	166	med	mono
comb	159	med	mono
stethoscope	16	lo	poly

Item	CELEX Lemma Frequency	Frequency Category	Syllable Length
hammock	19	lo	bi
swan	134	med	mono
magnet	52	lo	bi
sink	892	hi	mono
funnel	40	lo	bi
accordion	18	lo	poly
sling	63	lo	mono
microscope	135	med	poly
ball	1996	hi	mono
button	468	hi	bi
earring	59	lo	bi
desk	1633	hi	mono
nun	187	med	mono
banana	151	med	poly
handkerchief	351	hi	poly
dolphin	54	lo	bi
flag	461	hi	mono
calendar	151	med	poly
helicopter	281	hi	poly
snake	412	hi	mono
whistle	165	med	bi
buoy	12	hi	mono
zip	32	lo	mono
spider	126	med	bi
asparagus	38	lo	poly
octopus	27	lo	poly
pendulum	71	lo	poly
mop	49	lo	mono
tusk	33	lo	mono
handcuffs	34	lo	bi
peg	71	med	mono
triangle	131	med	poly
mushroom	227	hi	bi
vase	127	med	mono

Table B.3

Frequency and length of items in version 3 of the picture-naming task in order of appearance

Item	CELEX Lemma Frequency	Frequency Category	Syllable Length
castle	485	hi	bi
butcher	112	med	bi
tent	785	hi	mono
thermometer	116	med	poly
net	290	hi	mono
barrel	379	hi	bi
giraffe	28	lo	bi
cork	98	med	mono
furniture	696	hi	poly
butterfly	183	med	poly

Item	CELEX Lemma Frequency	Frequency Category	Syllable Length
scarf	219	hi	mono
pyjamas	146	med	poly
buckle	34	lo	bi
stool	222	hi	mono
glasses	571	hi	bi
trumpet	140	med	bi
microphone	152	med	poly
anchor	102	med	bi
tail	640	hi	mono
peacock	69	lo	bi
tambourine	13	lo	poly
spaghetti	82	med	poly
mirror	880	hi	bi
harp	50	lo	mono
escalator	30	lo	poly
zebra	34	lo	bi
corn	434	hi	mono
platypus	22	lo	poly
reflection	450	hi	poly
frog	168	med	mono
calculator	89	med	poly
dice	38	lo	mono
umbrella	245	hi	poly
star	1804	hi	mono
feather	379	hi	bi
skeleton	210	hi	poly
turtle	67	lo	bi
hose	72	lo	mono
computer	1683	hi	poly
newspaper	2176	hi	poly
canoe	101	med	bi
tripod	25	lo	bi
shadow	929	hi	bi
windmill	159	med	bi
rake	33	lo	mono
cherry	132	med	bi
Eskimo	31	lo	poly
saddle	177	med	bi
dart	57	lo	mono
pineapple	53	lo	poly
ostrich	48	lo	bi
snail	80	med	mono
television	2043	hi	poly
skunk	4	lo	mono
gorilla	54	lo	poly
plug	170	med	mono
kangaroo	48	lo	poly
safe	127	med	mono
kite	83	med	mono
hood	106	med	mono

Table B.4

Selection strength ratios and frequencies for items in version 1 of the verb generation task in order of appearance

Item	Selection Strength Ratio	Selection Strength Category	LEMMA Frequency Value	Frequency Category
barbeque	1.38	weak	39	lo
van	10	strong	1034	hi
crane	10	strong	71	lo
tail	2.17	weak	640	hi
razor	1.8	weak	156	lo
stethoscope	6.4	strong	16	lo
penny	2	weak	476	hi
curtains	2	weak	784	hi
baby	1.6	weak	4620	hi
heart	2.43	weak	2937	hi
ice	1.21	weak	944	hi
nun	41	strong	187	lo
bed	10.33	strong	4831	hi
shark	1.21	weak	357	lo
axe	1.05	weak	153	lo
yacht	18.5	strong	108	lo
mosquito	7.75	strong	96	lo
wool	11	strong	384	lo
piano	40	strong	488	hi
ladder	40	strong	287	lo
stomach	5.4	strong	769	hi
duck	2.17	weak	248	lo
hinge	2.67	weak	64	lo
ear	2.07	weak	1570	hi
worm	2.8	weak	302	lo
lion	39	strong	454	hi
boat	1.38	weak	1386	hi
fire	8	strong	2905	hi
ball	4.6	strong	1996	hi
pool	13.5	strong	733	hi
towel	40	strong	392	lo
trapeze	5	strong	9	lo
tongue	1.44	weak	715	hi
airplane	12.67	strong	102	lo
bell	11.67	strong	745	hi
fence	1.2	weak	537	hi
sparrow	1.5	weak	79	lo
kettle	19.5	strong	216	lo
sugar	6.75	strong	1015	hi
crab	1.11	weak	170	lo
chair	11.33	strong	2441	hi
picture	5.2	strong	3113	hi
feet	1.73	weak	5857	hi
caravan	1.8	weak	179	lo
leg	1.36	weak	3140	hi
radio	1.64	weak	1582	hi

Table B.5

Selection strength ratios and frequencies for items in version 2 of the verb generation task in order of appearance

Item	Selection Strength Ratio	Selection Strength Category	LEMMA Frequency Value	Frequency Category
elbow	17.5	strong	466	hi
church	10	strong	3287	hi
arrow	12	strong	264	lo
telephone	18	strong	1876	hi
scissors	100% response agreement	strong	79	lo
pills	1.21	weak	507	hi
stove	19.5	strong	364	lo
road	1.14	weak	4458	hi
daisy	1.23	weak	568	hi
watch	1.86	weak	710	hi
basket	13	strong	428	lo
key	1.5	weak	1544	hi
sun	9.67	strong	2728	hi
rope	5.2	strong	745	hi
pipe	20	strong	558	hi
alligator	2.78	weak	28	lo
shovel	100% response agreement	strong	76	lo
dice	6.8	strong	16	lo
binoculars	2.1	weak	9	lo
can	11	strong	166	lo
teeth	18	strong	56	lo
pan	2.78	strong	489	hi
hawk	5.25	strong	109	lo
scales	6.4	strong	1479	hi
frog	1	weak	8	lo
cigarette	2.89	weak	1274	hi
package	1.14	weak	357	lo
needle	1.83	weak	294	lo
tiger	2.5	weak	214	lo
horse	7.33	strong	2372	hi
straw	1.8	weak	461	hi
envelope	1.1	weak	83	lo
ghost	5.75	strong	554	hi
broom	13	strong	140	lo
seesaw	1.8	weak	12	lo
soldier	1.15	weak	1488	hi
candle	1.38	weak	294	lo
lips	1.78	weak	1401	hi
towel	40	strong	392	lo
moon	2.57	weak	1058	hi
priest	17.5	strong	873	hi
carnation	2.13	weak	28	lo
snow	6.5	strong	1102	hi
basin	9	strong	341	lo
suitcase	1.5	weak	334	lo

Table B.6

Frequencies and syllable length for version 1 of the picture-word verification task in order of appearance

Target	Distractor	Distractor Type	Frequency (Log CELEX lemma frequencies)	Syllable Length
scissors	curler	UR	0.78	2
cannon	cattle	Phon	2.83	2
bread	toast	Sem	3.23	1
scissors	scissors	Rel	2.53	2
cat	can	Phon	5.43	1
spider	spiral	Phon	1.96	2
cannon	cannon	Rel	2.65	2
spoon	file	UR	3.35	1
rabbit	cradle	UR	2.16	2
salad	satin	Phon	2.13	2
pencil	pencil	Rel	2.70	2
salad	vessel	UR	2.68	2
candle	whisker	UR	1.15	2
spoon	spear	Phon	2.37	1
drum	drip	Phon	2.42	1
pencil	chalk	Sem	2.26	1
cannon	pistol	Sem	2.71	2
drum	corn	UR	2.86	1
bread	bread	Rel	3.16	1
candle	torch	Sem	4.98	1
rabbit	rabbit	Rel	3.01	2
spider	gherkin	UR	0.70	2
scissors	dagger	Sem	4.92	2
pencil	pendant	Phon	1.67	2
pizza	peeler	Phon	1.20	2
spoon	mug	Sem	6.54	1
drum	bass	Sem	2.59	1
bread	form	UR	3.34	1
cannon	chin	UR	2.81	1
cat	cat	Rel	3.53	1
spider	spider	Rel	2.71	2
candle	canvas	Phon	2.34	2
rabbit	beaver	Sem	2.39	2
salad	pasta	Sem	2.34	2
pizza	burger	Sem	2.72	2
pencil	drama	UR	3.01	2
cat	frost	UR	2.39	1
bread	brain	Phon	3.59	1
pizza	gecko	UR	1.3	2
rabbit	rabbi	Phon	2.54	1
salad	salad	Rel	2.94	2
spoon	spoon	Rel	2.59	1
drum	drum	Rel	2.64	1
scissors	syrup	Phon	5.1	2
spider	cockroach	Sem	2.24	2
pizza	pizza	Rel	3.23	1

Target	Distractor	Distractor Type	Frequency (Log CELEX lemma frequencies)	Syllable Length
candle	candle	Rel	2.61	2
cat	lamb	Sem	2.73	1

Table B.7

Frequencies and syllable length for version 2 of the picture-word verification task in order of appearance

Target	Distractor	Distractor Type	Frequency (Log CELEX lemma frequencies	Syllable Length
trumpet	violin	Sem	2.39	2
grapes	grease	Phon	2.55	1
hammer	pearl	UR	2.90	1
knife	limb	UR	2.38	1
trumpet	truffle	Phon	1.38	2
horse	deer	Sem	2.65	1
chair	stool	Sem	2.26	1
lemon	leather	Phon	2.84	2
grapes	puzzle	UR	2.57	2
razor	perfume	UR	2.77	2
turkey	turkey	Rel	3.06	2
knife	fork	Sem	2.65	1
carrot	spinach	Sem	2.12	2
monkey	monkey	Rel	3.23	2
lemon	lemon	Rel	2.79	2
trumpet	denim	UR	1.53	2
carrot	carrot	Rel	2.29	2
trumpet	trumpet	Rel	2.32	2
horse	linen	UR	2.18	2
arrow	torch	UR	2.41	1
chair	chair	Rel	3.40	1
razor	radar	Phon	3.26	2
monkey	emerald	UR	2.12	3
arrow	Arab	Phon	2.24	1
chair	nickel	UR	2.64	2
turkey	star	UR	3.62	1
grapes	peach	Sem	2.51	1
horse	horse	Rel	3.68	1
lemon	scarf	UR	2.38	1
monkey	panda	Sem	2.04	2
razor	razor	Rel	2.54	2
arrow	arrow	Rel	2.60	2
hammer	axe	Sem	2.4	1
chair	check	Phon	4.15	1
monkey	mustard	Phon	2.52	2
turkey	duck	Sem	3.1	1
knife	knife	Rel	3.38	1
turkey	turban	Phon	1.83	2
lemon	orange	Sem	3.06	2

Target	Distractor	Distractor Type	Frequency (Log CELEX lemma frequencies	Syllable Length
knife	nine	Phon	3.54	1
grapes	grapes	Rel	2.31	1
hammer	hamlet	Phon	2.37	2
arrow	pipe	UR	3.00	1
carrot	chasm	Phon	1.28	1
horse	haunt	Phon	2.26	1
hammer	hammer	Rel	2.80	2

Table B.8

Frequencies and imageability of the single word repetition task in order of appearance

Item	Freq_KF	LogFreq_KF	Freq Category	Imageability	Imageability Category
episode	12	1.08	Lo	370	Lo
theory	129	2.11	Hi	317	Lo
potato	15	1.18	Lo	617	Hi
church	348	2.54	Hi	616	Hi
folly	10	1.00	Lo	326	Lo
irony	12	1.08	Lo	293	Lo
battle	87	1.94	Hi	597	Hi
concept	85	1.93	Hi	258	Lo
spider	2	0.30	Lo	597	Hi
village	72	1.86	Hi	578	Hi
deed	8	0.90	Lo	390	Lo
gravy	4	0.60	Lo	594	Hi
dogma	4	0.60	Lo	327	Lo
alcohol	13	1.11	Lo	598	Hi
picture	162	2.21	Hi	581	Hi
radio	120	2.08	Hi	613	Hi
onion	15	1.18	Lo	617	Hi
purpose	149	2.17	Hi	280	Lo
quality	114	2.06	Hi	349	Lo
school	492	2.69	Hi	599	Hi
system	416	2.62	Hi	340	Lo
elephant	7	0.85	Lo	616	Hi
pig	8	0.90	Lo	635	Hi
night	411	2.61	Hi	607	Hi
marriage	95	1.98	Hi	556	Hi
thing	333	2.52	Hi	358	Lo
bonus	2	0.30	Lo	397	Lo
opinion	96	1.98	Hi	359	Lo

Item	Freq_KF	LogFreq_KF	Freq Category	Imageability	Imageability Category
analogy	13	1.11	Lo	267	Lo
hand	431	2.63	Hi	598	Hi
woe	5	0.70	Lo	348	Lo
character	118	2.07	Hi	372	Lo
wheat	9	0.95	Lo	577	Hi
effort	145	2.16	Hi	367	Lo
tribute	24	1.38	Lo	386	Lo
fact	447	2.65	Hi	302	Lo
valour	N/A	-	Lo	-	Lo
idea	195	2.29	Hi	319	Lo
axe	6	0.78	Hi	597	Hi
funnel	1	0.00	Lo	-	Hi
tractor	24	1.38	Lo	585	Hi
length	116	2.06	Hi	395	Lo
plea	11	1.04	Lo	347	Lo
monkey	9	0.95	Lo	588	Hi
manner	124	2.09	Hi	342	Lo
satire	9	0.95	Lo	370	Lo
drum	11	1.04	Lo	599	Hi
cart	5	0.70	Lo	597	Hi
miracle	16	1.20	Lo	367	Lo
hospital	110	2.04	Hi	60	Hi
audience	115	2.06	Hi	555	Hi
attitude	8	0.90	Lo	321	Lo
letter	145	2.16	Hi	595	Hi
tobacco	19	1.28	Lo	601	Hi
principle	109	2.04	Hi	305	Lo
plane	114	2.06	Hi	556	Hi
moment	246	2.39	Hi	334	Lo
summer	134	2.13	Hi	618	Hi
feather	6	0.78	Lo	-	Hi
pact	5	0.70	Lo	364	Lo

Table B.9

List of items in the nonword repetition task and their associated actual word, in order of appearance.

Nonword	Associated Word
biffle	battle
ragio	radio
clee	deed
otion	onion
drim	drum

sping	thing
slurch	student
plen	plea
atalogy	analogy
parpise	purpose
loment	moment
lutter	letter
hend	hand
trantor	tractor
voe	woe
merly	mercy
baranter	character
affort	effort
gramy	gravy
spunder	spider

Table B.10

Items, their associated colours and condition in the Stroop task in order of appearance

Item	Colour	Condition
pink	pink	Congruent
brown	grey	Incongruent
red	red	Congruent
purple	yellow	Incongruent
green	orange	Incongruent
blue	purple	Incongruent
orange	orange	Congruent
yellow	blue	Incongruent
purple	red	Incongruent
blue	blue	Congruent
red	green	Incongruent
green	yellow	Incongruent
blue	green	Incongruent
orange	purple	Incongruent
yellow	yellow	Congruent
red	blue	Incongruent
purple	purple	Congruent
orange	red	Incongruent
green	green	Congruent
yellow	orange	Incongruent

Table B.11

Items presented in order of appearance in the Articulatory Agility test

Item
mamma
tip-top
fifty-fifty
thanks
huckleberry
baseball player
caterpillar

Table B.12

Items Presented in Order of Appearance for the Nonword Reading Task

Item
ked
nar
fon
shid
doop
dusp
snite
hoach
glope
dringe
churse
shoave

Appendix C:

Neurological tumour patients: Procedure and Materials

Subtest Administration Order

Table C.1.

Task administration order during the follow-up testing phase.

Test	Format	Chapter
Simple Reaction Time	Computer	4
Design Fluency	Pen and paper	4
Choice Reaction Time	Computer	4
QPA	Oral	3
Prepare Reaction Time	Computer	4
TROG	Oral and pen and paper	3
Concentrate	Computer	4
Antisaccade	Computer	4
BLAST subtests		4

Information and Consent Form for Brain Tumour Patients

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA
UNIVERSITY OF WELLINGTON

Information Sheet
Study: The assessment of language before, and after neurosurgery

Katie Fowler	Dr Carolyn Wilshire, PhD	Kay Cunningham, MA, DipClinPsyc
PhD Candidate	**Primary Supervisor**	**Clinical Supervisor**
Katherine.Fowler@vuw.ac.nz	Carolyn.Wilshire@vuw.ac.nz	Kay.Cunningham@xtra.co.nz
	(04) 463 6036	

You are invited to take part in a research project for Victoria University of Wellington. Please take your time to read through the information sheet. Your participation is entirely voluntary (your choice). You do not have to take part in this study, and if you choose not to take part you will receive the standard treatment/care available. Participation in this study will be stopped should any harmful effect appear or if the doctor feels it is not in your best interest to continue

What is the purpose of this research?
- This research will allow us to assess the effectiveness of various different kinds of language tasks for assessing language before, and after neurosurgery for the removal of a brain tumor. The results obtained may help clinicians to assess language more effectively in this context. The information we gain may also provide useful insights into human language more specifically, and the way that it is organised in the brain.

Who is selected for the study?
- All patients undertaking neurosurgery for the removal of a brain tumor in Wellington hospital are invited to participate in this study
-

Where will the study take place?
- Testing before and after will take place at the Neurology department in Wellington hospital. Participants may also be asked if they would be willing to participate in a follow-up visit at their homes 6-8 weeks following surgery.

What is involved if you agree to participate?
- If you agree to participate in this study you will partake in two testing sessions; before surgery, and after surgery. During each session, you'll be asked to do a range of simple language tasks that aim to test a range of language abilities.
- The language tests will involve you naming pictures, giving a list of words that start with a particular letter of the alphabet and category, repeating words presented by the examiner, giving an action word that is associated with an object and reading coloured words whilst ignoring their colour.

- For the before- and after-surgery testing, we will visit you in your hospital ward at a time that suits you.
- We anticipate that your total involvement will take no more than one hour per session.
- During the research, you are free to withdraw at any point before your data has been collected. You may participate in any or all testing phases, it's up to you.

Who is conducting the research?
- Katie Fowler is a PhD student at Victoria University and is undertaking this research as part of her doctorate thesis. Ms Cunningham, is a registered Clinical Neuropsychologist. Dr. Wilshire, a Senior Lecturer at the School of Psychology at Victoria University of Wellington, is the primary supervisor for this project.

Privacy and Confidentiality
- We will keep your consent forms and data until the research project has been completed and the findings are published.
- You will never be identified in this research project or in any other presentation or publication. The information you provide will be coded by number or initials.
- In accordance with the requirements of some scientific journals and organizations, your coded data may be shared with other competent researchers.
- Your coded data may be used in other, related studies.
- A copy of the coded data will remain in the custody of Dr. Wilshire.

What happens to the information that you provide?
- The data you provide may be used for one or more of the following purposes:
  • The overall findings may be submitted for publication in a scientific journal, or presented at scientific conferences.
  • The overall findings will form part of PhD thesis, that will be submitted for assessment.

What are the benefits of this study?
- This study will allow participants to obtain detailed feedback about their specific strengths and weaknesses
- Given that the language evaluation will be more extensive than is normally given in these cases, the information we gain may be more useful for other health practitioners

What are the risks of this study?
- Participants may be experiencing fatigue and possible distress during testing before and after brain surgery. To minimize any potential harm, testing sessions are kept as short as possible and remember, you are free to stop at any time for any reason.

What are the inclusion and exclusion criteria for this study?
- People invited to participate in this study will be those who:
  - Have been admitted to Wellington Hospital for neurosurgery
  - Are at least 18 years of age
  - Whose surgery will focus on a key brain region involved in language
- Those not eligible are:
  - Those whose native language is not English
  - Anyone who, in the opinion of the individual's surgical team, may find the testing unduly stressful.

Results

- If you would like to know the results of this study, they will be available approximately in December 2017 from the following sources:
  - Information posted/emailed to you upon request

Statement of Approval

- This study has received ethical approval from The New Zealand Health and Disability Ethics Committee, ethics reference number CEN/11/07/037

If you have any queries or concerns regarding your rights as a participant in this study, you may wish to contact an independent health and disability advocate:

Free phone: 0800 555 050
Free fax: 0800 2 SUPPORT (0800 2787 7678)
Email: advocacy@hdc.org.nz

If you have any further questions regarding this study, please contact any one of us above.

VICTORIA
UNIVERSITY OF WELLINGTON

Consent Form
Study: The assessment of language before and after neurosurgery.

I have read and I understand the information sheeted dated _____ for volunteers taking part in the study designed to test specific language functioning before and after neurosurgery.

I have had the opportunity use whānau support or a friend to help me ask questions and understand the study.

I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time, and this will in no way affect my future health care and academic progress.

I understand that my participation in this study is confidential and that no material that could identify me will be used in any reports on this study.

I have had time to consider whether to take part in the study

I know whom to contact if I have side-effects from the study

I know whom to contact if I have any questions about the study in general

I consent to my interview being audiotaped   Yes ☐   No ☐

I wish to receive a copy of the results…………………….   Yes ☐   No ☐

I _____ hereby consent to take part in this study

Date:

Signature:

Full names of researchers:

Contact phone number for researchers:

Project explained by:

Project role:

Signature:

Date:

VICTORIA
UNIVERSITY OF WELLINGTON

Letter for Follow-up Patients

Name
Street Address
Suburb
City   Date

Dear Mr/Mrs,

During your recent visit to Wellington Hospital you took part in a language study being conducted by Katie Fowler and Dr. Carolyn Wilshire from Victoria University. This study investigated the language skills of people about to undergo surgery for a tumour, and how these change after the surgery.

I'm writing now to ask if you would be willing to participate in a follow-up study. This new study will explore whether people's language skills continue to improve after they have fully recovered from their surgery. You would be asked to do the same language tests you did before, plus several new ones.

No travel is required; we are happy to visit you in your home or any other location you'd like, at a date and time that suits you. Overall, this follow-up study is expected to take between one to one and a half hours, however this can be split into two sessions if you prefer. However, the choice to participate is yours – you're under no obligation to take part if you prefer not to.

I will be in contact via phone in approximately one week to see whether you would be interested; however, if you have any questions please do not hesitate to contact me at Katherine.Fowler@vuw.ac.nz

Kind regards,

Ms. Katherine Fowler   Dr. Carolyn Wilshire
Graduate researcher in Neuropsychology   Senior Lecturer in Neuropsychology
Email Katherine.fowler@vuw.ac.nz   Email Carolyn.Wilshire@vuw.ac.nz
                                             Tel (04) 4626036

BLAST Tables

BLAST Individual Subtests

Table D.1.

Individual T scores for each patient for each BLAST subtest. Red cells represent significant impairment – bold and italicised cells reflect T scores less than 30 and significant impairment relative to the appropriate age-matched control group using Crawford and Howell's (1998) modified t test, whilst red cells that are not bolded and italicised reflect T scores less than 30, but not significantly impaired according to Crawford and Howell's (1998) modified t test. N.B. Patients SN, DP, and WR did not complete pre or post-operative testing

	Picture Naming			Verb Generation			Picture-Word Verification		
	Pre	**Post**	**Follow-up**	**Pre**	**Post**	**Follow-up**	**Pre**	**Post**	**Follow-up**
Left Frontal									
BP	52.49	45.40	27.59	41.43	44.00	44.03	61.68	56.20	53.36
RF	52.65	51.64	36.33	*5.46*	*25.15*	38.85	*1.43*	53.89	53.89
TD	46.64	na	60.96	39.95	na	44.41	36.81	na	53.36
LC	52.72	*3.50*	42.67	44.40	*17.05*	48.84	36.81	*-10.31*	*21.53*
Average	**51.13**	**33.51**	**41.89**	**32.81**	**28.73**	**44.03**	**34.18**	**33.26**	**45.53**
SD	**2.99**	**26.18**	**14.14**	**18.33**	**13.83**	**4.09**	**24.78**	**37.75**	**16.01**
Left Temporal									
CA	*-44.96*	*16.65*	30.49	*-31.34*	*20.00*	*8.92*	*-61.36*	53.36	53.36
SH	*19.23*	*22.46*	*-106.53*	27.74	37.67	*16.56*	*-14.59*	*-3.98*	*15.28*
GM	*14.42*	*-21.29*	*-5.99*	31.44	*29.32*	37.86	*-61.20*	*-34.01*	*-23.33*
KW	*38.26*	*17.36*	34.71	54.40	53.40	54.42	*24.27*	56.20	53.36

Average	**6.74**	**8.80**	**-11.83**	**20.56**	**35.10**	**29.44**	**-28.22**	**17.89**	**24.67**
SD	**35.97**	**20.22**	**65.72**	**36.55**	**14.18**	**20.67**	**41.34**	**44.34**	**36.69**
Right Frontal									
NH	52.72	29.81	48.78	*22.25*	31.81	*15.57*		na	37.40
DF	40.52	47.36	48.78	31.10	40.67	42.20	*25.92*	54.07	37.40
BS	49.70	44.33	46.93	61.13	29.32	38.85	32.78	53.89	53.89
LW	49.70	55.27	9.88	20.31	*-8.26*	42.56		*-3.98*	34.54
GP	*14.42*	*11.52*	*-48.31*	38.86	*16.79*	*-27.97*	*1.43*	34.54	53.89
RS	40.89	44.33	41.64	49.99	62.74	25.72	64.14	56.20	53.89
Average	**41.33**	**38.77**	**24.62**	**37.27**	**28.85**	**22.82**	**24.85**	**38.94**	**45.17**
SD	**14.11**	**15.69**	**38.71**	**16.03**	**23.75**	**27.06**	**26.33**	**25.55**	**9.61**
Left Parietal									
AVG	58.84	60.52	48.78	48.83	43.62	62.16	36.81	37.40	37.40
BD	37.93	*22.46*	*25.76*	35.15	50.20	57.42	48.42	53.89	*-5.28*
CR	32.06	51.64	52.20	35.15	29.32	54.04	64.14	*-3.98*	53.89
NOH	32.06	na	31.06	46.29	na	50.00	*-76.84*	na	*-3.98*
KG	58.84	56.12	36.56	35.53	46.57	48.84	58.64	53.36	53.36
SO	29.11	*22.46*	*-5.99*	*5.46*	*-4.09*	*-2.02*	*17.14*	34.63	34.54
Average	**41.48**	**42.64**	**31.39**	**34.40**	**33.13**	**45.07**	**24.72**	**35.06**	**28.32**
SD	**13.75**	**18.69**	**20.93**	**15.42**	**22.26**	**23.59**	**52.51**	**23.55**	**26.74**
Right Posterior									
SG	34.40	38.60	30.49	*17.80*	34.77	22.24	*25.92*	*21.53*	37.40
AEK	35.01	*15.18*	15.19	38.86	*4.26*	61.15	48.42	53.89	*-3.98*
JB	40.52	56.12	48.78	*-1.32*	*20.00*	53.27	47.70	53.36	53.36

PC	na	*12.29*	30.49	35.53	43.62	35.61	na	*5.65*	37.40
Average	**36.64**	**30.55**	**31.24**	**22.72**	**25.66**	**43.07**	**40.68**	**33.61**	**31.05**
SD	**3.37**	**20.73**	**13.74**	**18.50**	**17.28**	**17.52**	**12.79**	**24.01**	**24.53**
AE	*3.89*	*-14.03*	*-6.10*	*8.95*	*2.29*	53.27	*-6.81*	*5.65*	53.36

	Letter Fluency			Category Fluency			Congruent Stroop		
	Pre	**Post**	**Follow-Up**	**Pre**	**Post**	**Follow-Up**	**Pre**	**Post**	**Follow-Up**
Left Frontal									
BP	28.40	41.49	30.59	45.94	47.07	45.94	50.00	50.00	52.36
RF	37.54	38.25	35.38	29.60	39.16	36.77	52.27	52.27	52.27
TD	*23.90*	na	27.74	29.31	na	33.37	50.00	na	50.00
LC	38.28	*19.11*	38.28	48.28	*10.34*	45.57	50.00	50.00	50.00
Average	**32.03**	**32.95**	**33.00**	**38.28**	**32.19**	**40.41**	**50.57**	**50.76**	**51.16**
SD	**7.04**	**12.10**	**4.73**	**10.24**	**19.33**	**6.32**	**1.14**	**1.31**	**1.34**
Left Temporal									
CA	*20.07*	31.57	44.04	*3.56*	*17.11*	*21.18*	50.00	50.00	50.00
SH	33.94	37.54	*24.60*	37.57	33.59	29.60	na	na	6.82
GM	30.35	33.94	35.38	34.38	29.60	38.37	*-38.64*	*6.82*	52.27
KW	na	*28.40*	30.59	na	*26.76*	38.04	na	50.00	52.36
Average	**28.12**	**32.86**	**33.65**	**25.17**	**26.77**	**31.80**	**5.68**	**35.61**	**40.36**
SD	**7.20**	**3.85**	**8.21**	**18.78**	**7.02**	**8.16**	**62.68**	**24.93**	**22.39**
Right Frontal									
NH	44.04	44.04	34.45	49.63	50.99	52.34	50.00	50.00	50.00
DF	*21.98*	*22.94*	33.49	33.37	*25.24*	34.73	50.00	50.00	na

BS	46.88	na	41.13	53.51	na	54.30	52.27	52.27	50.00
LW	30.35	30.35	36.82	44.74	35.98	46.33	52.27	52.27	52.27
GP	31.79	39.69	36.10	37.57	43.15	39.16	*6.82*	*6.82*	*6.82*
RS	49.03	48.31	51.19	46.33	39.96	48.73	52.27	52.27	52.27
Average	**37.34**	**37.07**	**38.86**	**44.19**	**39.06**	**45.93**	**43.94**	**43.94**	**42.27**
SD	**10.84**	**10.33**	**6.59**	**7.51**	**9.49**	**7.62**	**18.22**	**18.22**	**19.85**
Left Parietal									
AVG	*26.78*	*23.90*	59.38	40.15	48.28	45.57	50.00	50.00	50.00
BD	*28.20*	*26.76*	36.10	50.32	45.54	59.88	52.27	52.27	52.27
CR	42.56	na	62.68	37.57	na	48.73	*6.82*	na	52.27
NOH	*26.04*	*25.32*	*28.20*	29.60	*28.81*	35.18	52.27	52.27	50.00
KG	40.20	33.49	32.53	46.92	34.73	34.73	na	50.00	50.00
SO	33.23	44.72	39.69	37.57	41.55	32.79	52.27	52.27	52.27
Average	**32.83**	**30.84**	**43.10**	**40.36**	**39.78**	**42.81**	**42.73**	**51.36**	**51.14**
SD	**7.12**	**8.59**	**14.44**	**7.40**	**7.98**	**10.56**	**20.10**	**1.24**	**1.24**
Right Posterior									
SG	34.45	34.45	46.91	32.02	34.73	40.15	50.00	50.00	50.00
AEK	39.69	38.97	31.79	45.54	47.13	47.13	52.27	52.27	52.27
JB	47.87	46.91	44.04	48.28	49.63	59.12	50.00	50.00	50.00
PC	38.28	31.57	34.45	36.08	36.08	40.15	50.00	50.00	50.00
Average	**40.07**	**37.98**	**39.30**	**40.48**	**41.89**	**46.64**	**50.57**	**50.57**	**50.57**
SD	**5.65**	**6.69**	**7.31**	**7.69**	**7.58**	**8.95**	**1.14**	**1.14**	**1.14**
AE	33.49	33.49	*24.86*	37.44	*17.11*	*25.24*	50.00	50.00	50.00

Incongruent Stroop **Articulation**

	Pre	Post	Follow-up	Pre	Post	Follow-up
Left Frontal						
BP	54.47	54.47	80.79	35.23	45.83	55.36
RF	56.71	56.71	50.00	45.05	43.54	47.32
TD	54.47	na	28.16	2.39	na	40.53
LC	54.47	54.47	80.79	35.23	35.23	41.59
Average	**55.03**	**55.22**	**59.93**	**29.48**	**31.15**	**46.20**
SD	**1.12**	**1.29**	**25.68**	**18.64**	**21.26**	**6.80**
Left Temporal						
CA	*1.84*	54.47	80.79	32.06	24.64	23.58
SH	NA	NA	56.71	42.03	41.28	42.03
GM	*16.44*	36.58	*23.15*	34.48	33.72	47.32
KW	NA	80.79	80.79	53.24	61.72	54.30
Average	**9.14**	**57.28**	**60.36**	**40.45**	**40.34**	**41.81**
SD	**10.32**	**22.24**	**27.28**	**9.53**	**15.79**	**13.15**
Right Frontal						
NH	54.47	54.47	80.79	na	na	40.53
DF	80.79	54.47	na	1.33	34.17	na
BS	56.71	56.71	43.29	56.38	na	57.14
LW	50.00	43.29	50.00	48.83	51.10	52.61
GP	*23.15*	*23.15*	29.87	40.52	47.32	46.56
RS	56.71	56.71	56.71		51.85	54.12
Average	**53.64**	**48.14**	**52.13**	**36.77**	**46.11**	**41.83**
SD	**18.44**	**13.23**	**18.85**	**24.49**	**8.20**	**21.33**
Left Parietal						

AVG	54.47	54.47	*1.84*	68.07	69.13	53.24
BD	50.00	43.29	36.58	51.10	56.38	56.38
CR	43.29	NA	63.42	39.01		42.03
NOH	29.87	29.87	56.71	42.03	35.23	51.85
KG	NA	28.16	*1.84*	56.42	62.78	45.83
SO	50.00	29.87	50.00	31.46	29.19	39.01
Average	**45.53**	**37.13**	**35.07**	**48.01**	**50.54**	**48.06**
SD	**9.62**	**11.45**	**27.22**	**13.22**	**17.46**	**6.84**
Right Posterior						
SG	54.47	*-24.47*	*1.84*	47.94	51.12	51.12
AEK	56.71	56.71	56.71	51.10	42.03	45.81
JB	54.47	54.47	80.79	64.89	61.72	37.35
PC	80.79	*1.84*	*1.84*	70.19	74.43	47.94
Average	**61.61**	**22.14**	**35.30**	**58.53**	**57.32**	**45.56**
SD	**12.83**	**40.11**	**39.86**	**10.71**	**13.95**	**5.89**
AE	*-24.47*	*-50.79*	*-50.79*	36.29	42.65	35.23

Table D.2.

Individual T scores for each patient for each BLAST core skill. Red cells represent a significant impairment relative to Crawford and Howell's (1998) modified t test.

Patient	Auditory Word Recognition		Accessing Semantic Knowledge		Lexical Selection		Phonological Encoding	
	Pre	Follow-up	Pre	Follow-up	Pre	Follow-up	Pre	Follow-up
Left Frontal								
BP	53.12	53.25	67.07	63.87	39.25	34.68	50.42	42.29
LC	53.75	52.18	50.04	40.58	52.73	42.71	41.96	52.98
TD	53.75	54.68	50.21	22.19	70.50	69.48	na	46.77
RF	43.77	56.89	30.56	46.66	57.74	48.31	63.48	55.62
Average	**51.10**	**54.25**	**49.47**	**43.32**	**55.05**	**48.80**	**51.95**	**49.41**
SD	**4.89**	**2.04**	**14.93**	**17.20**	**12.92**	**14.88**	**10.84**	**6.03**
Left Parietal								
CR	26.10	56.89	42.19	36.00	62.15	53.04	37.25	48.24
NOH	17.74	47.79	33.84	50.78	36.51	41.03	51.24	57.91
SO	52.98	56.89	33.33	30.49	61.54	27.49	51.99	47.01
AVG	53.75	54.68	28.85	37.34	75.42	66.88	43.77	42.03
BD	52.98	42.33	38.02	42.35	53.79	46.34	61.83	47.30
KG	53.75	54.68	56.85	49.39	55.83	50.97	64.11	62.51

Average	**42.88**	**52.21**	**38.84**	**41.06**	**57.54**	**47.63**	**51.70**	**50.83**
SD	**16.45**	**5.88**	**9.91**	**7.96**	**12.78**	**13.13**	**10.28**	**7.72**
Left Temporal								
CA	37.95	54.68	19.89	25.28	75.91	46.85	51.78	57.71
KW	na	53.25	30.62	50.20	na	27.58	15.27	-3.13
SH	na	56.89	41.83	44.24	-7.83	11.24	65.50	49.63
GM	na	52.40	28.97	27.55	25.31	47.62	31.91	37.71
Average	**37.95**	**54.31**	**30.33**	**36.82**	**31.13**	**33.32**	**41.11**	**35.48**
SD		**1.96**	**9.00**	**12.29**	**42.17**	**17.40**	**22.07**	**27.02**
Right Frontal								
DF	37.95	54.68	51.20	45.91	58.68	50.90	61.77	na
LW	52.98	60.80	61.61	53.35	42.60	55.04	46.81	35.85
BS	43.77	56.89	51.89	64.71	54.78	37.01	49.04	47.79
GP	10.82	56.89	43.91	44.07	53.95	31.44	42.54	36.59
NH	na	42.18	na	63.69	na	42.75	na	47.30
RS	52.98	56.89	44.79	47.86	53.57	49.94	na	54.24
Average	39.70	54.72	50.68	53.27	52.72	44.51	50.04	44.35
SD	17.37	6.46	7.10	9.03	6.01	9.07	8.27	7.92
Right Posterior								
PC	na	54.68	na	52.66	na	71.75	na	50.15

Initials								
JB	53.75	54.68	47.57	61.65	41.50	38.26	52.40	57.62
AEK	47.85	56.89	55.37	57.47	55.58	50.67	52.11	42.42
SG	na	54.68	36.62	38.14	76.31	51.45	na	55.20
Average	**51.52**	**55.56**	**46.09**	**51.56**	**56.74**	**52.42**	**52.25**	**51.93**
SD	**3.21**	**1.21**	**7.74**	**9.11**	**14.45**	**12.08**	**0.20**	**5.96**
AE	34.79	36.43	45.40	43.25	31.17	27.34	65.76	57.05

	Verbal Short-term Memory		Verb Retrieval		Articulation		Goal-Driven Response Selection	
Initials	Pre	Follow-up	Pre	Follow-up	Pre	Follow-up	Pre	Follow-up
Left Frontal								
BP	-3.33	36.67	50.60	39.89	22.43	55.79	27.60	47.46
LC	41.24	42.12	56.94	39.78	34.33	41.58	42.59	44.99
TD	41.24	56.31	44.84	13.72	na	16.36	43.78	58.46
RF	56.66	59.61	10.81	45.15	39.44	47.32	42.34	50.13
Average	**33.95**	**48.67**	**40.80**	**34.63**	**32.07**	**40.26**	**39.08**	**50.26**
SD	**25.90**	**11.03**	**20.59**	**14.17**	**8.73**	**16.97**	**7.68**	**5.86**
Left Parietal								
CR	36.05	41.94	66.79	60.09	34.46	42.04	49.27	60.62
NOH	50.77	59.61	51.25	36.91	45.84	51.85	20.08	63.51

SO	59.61	53.72	23.22	9.85	31.62	39.02	53.26	48.26
AVG	55.57	na	52.86	61.35	65.27	53.23	46.50	65.41
BD	59.61	50.77	42.52	56.68	50.11	56.38	50.22	58.06
KG	55.57	56.31	48.21	61.35	35.32	45.82	na	37.29
Average	**52.86**	**52.47**	**47.48**	**47.70**	**43.77**	**48.06**	**43.87**	**55.53**
SD	**8.86**	**6.72**	**14.35**	**20.77**	**12.74**	**6.83**	**13.51**	**10.76**
Left Temporal								
CA	41.24	13.74	-13.97	11.84	31.33	23.56	53.66	45.32
KW	na	16.67	54.89	59.86	50.38	54.38	34.69	50.54
SH	na	59.61	43.16	-0.72	41.57	42.04	na	31.78
GM	na	47.83	36.92	81.75	34.46	33.73	44.03	65.71
Average	**41.24**	**34.46**	**30.25**	**38.18**	**39.44**	**38.43**	**44.13**	**48.34**
SD	**-**	**22.78**	**30.40**	**39.05**	**8.46**	**13.04**	**9.48**	**14.02**
Right Frontal								
DF	55.57	56.31	42.16	46.72	41.31	na	34.59	na
LW	50.77	47.83	51.25	62.10	47.97	52.60	32.79	45.33
BS	59.61	50.77	53.73	24.79	55.08	57.13	52.27	51.24
GP	53.72	53.72	60.54	-1.32	40.15	39.77	37.23	52.49
NH	na	13.74	10.63	-1.70	34.33	40.52	41.87	31.54
RS	59.61	59.61	72.39	45.15	na	54.11	51.52	59.12
Average	**54.92**	**44.47**	**43.66**	**26.12**	**43.77**	**47.51**	**39.75**	**45.15**
SD	**3.70**	**17.47**	**19.61**	**28.49**	**7.97**	**8.70**	**7.79**	**9.60**

Right Posterior							
PC	na	56.31	na	54.31	67.26	47.94	na	60.13
JB	na	27.93	52.86	48.04	62.27	37.34	45.83	47.78
AEK	56.66	53.72	66.79	70.92	50.11	45.81	44.70	44.80
SG	na	56.31	18.66	15.03	46.30	51.12	45.49	51.95
Average	**58.13**	**50.77**	**52.68**	**46.69**	**56.49**	**47.26**	**46.89**	**52.76**
SD	**2.08**	**12.94**	**24.12**	**20.32**	**9.90**	**6.38**	**3.13**	**6.77**
AE	41.24	56.31	27.39	55.63	35.32	35.22	52.44	47.16

Healthy control participants' method and procedure

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA
UNIVERSITY OF WELLINGTON

Information Sheet:

Comparing controls on tasks measuring attention and speed

Katie Fowler   Dr Carolyn Wilshire
PhD Student   Primary Supervisor
Katherine.fowler@vuw.ac.nz   Carolyn.wilshire@vuw.ac.nz
04 463 5398 ext 7099   04 463 6036

What is the purpose of this research?
- The aim of this research is to learn more about how brain tumours can affect language and related cognitive functions. To achieve this, we have developed a set of tasks to assess these skills in tumour patients three months after their surgery. But before we can use these, we first need to find out how normal, healthy participants perform on them. We need people of a comparable age to our participants with neurological tumours.

Who is conducting the research?
- We are researchers in the School of Psychology at Victoria University of Wellington. Katie Fowler will be conducting the testing sessions. Dr. Wilshire is supervising the project. This research has been approved by the School of Psychology Human Ethics Committee under delegated authority of Victoria University of Wellington's Human Ethics Committee.

What is involved if you agree to participate?
- If you agree to participate, you will take part in one testing session of approximately 30 minutes' duration. During the session, you will be asked to do a range of simple tasks that aim to test a range of language and cognitive abilities.
- The tests may involve you pressing a button/s on the keyboard each time you see a certain letter or word appear onscreen, or seeing an arrow and reporting its direction, repeating a series of digits, and counting tones and numbers.
- The session will be audio recorded, but these recordings will be destroyed as soon as they have been transcribed.
- *(For IPRP participants only)* You will receive .5 IPRP credit for participating in this research.
- *(For non-IPRP participants)* You will receive $10 cash as reimbursement for your time, and a further $10 contribution towards travel expenses if you travel to the University to participate.

What happens to the information that you provide?

- Your audio recordings will not be stored nor shared in any way: they will be destroyed as soon as we have transcribed them.
- We will keep your consent forms and email addresses for up to 5 years after our findings have been published.
- We will keep copies of your coded and digital data indefinitely.
- Your data will be stored on the School of Psychology computer server, and the principal investigator's hard drive; access is password-protected and restricted only to investigators in Dr Wilshire's laboratory,
- Your coded data may be used in other, related studies or shared with other competent researchers within the laboratory.
- The data you provide may appear in theses, scientific publications or book chapters, and may be presented at scientific conferences. However, you will never be identified in any of these.
-

If you would like us to email you a copy of the results of this study, you can indicate this on the consent form. They will be available approximately in April 2016

If you have any further questions regarding this study please contact any one of us above.

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA
UNIVERSITY OF WELLINGTON

Consent to Participate in Research

Principal Investigator: Katie Fowler, School of Psychology

I have been given an explanation of the research project, and I understood what is involved. I have been given the opportunity to ask questions and have them answered to my satisfaction. I understand that I may withdraw myself (or any information I have provided) from this project at any time up until the data has been collected without having to give reasons.

I understand that any information that identifies me personally will be kept confidential to the investigators. I understand that the results of the study may be published in academic journals, presented in conference presentations, or shared with other competent professionals, but that any such information will not use my name, or any information that could identify me.

I understand that the session I take part in may be audio taped.

I understand that the data will be stored in a locked cabinet and/or a password-protected computer in Dr. Wilshire's research laboratory for fifteen years following publication, and then destroyed.

I .... agree to take part in this research

Signature ....

Date ....

I WOULD/WOULD NOT like to receive a summary of the results of the experiment when it is completed. *(Circle your choice).*

If yes, please provide your email address:

Email   _____

Neurological Status Questionnaire:

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA
UNIVERSITY OF WELLINGTON

Neurological Status Questionnaire
Study: A New Test Battery for Examining Language in Individuals Undergoing Neurosurgery

Principal Investigator: Katie Fowler, School of Psychology
It would help us with our research if you were able to provide the following additional information. Please note that completion of the following questions is entirely optional and confidential.

Age: Sex: M/F   Handedness: L/R
Highest level of education obtained: _____

_____

Have you ever suffered from any visual and/or hearing impairments? (please specify)

_____

_____

_____

_____

Have you ever experienced a neurological event such as a stroke, or other brain injury? (please specify)

_____

_____

_____

_____

_____

_____

_____

Cognitive Control Results

Table F.1

Each patient's z-scores on the individual component measures of our selected cognitive control skills. Cells highlighted in blue represent significant impairment according to Crawford and Howell's (1998) modified t test. Cells highlighted in grey (performance monitoring measures) represent those scores that we were unable to statistically compare.

	Attentional Maintenance Measures				Activation Measures		
	Average scores ISI	Prepare 1 vs 3	SRT - CRT	Slope	Average RT	Antisaccade RT	Congruent
Left Frontal							
BP	-0.194	0.476	-0.373	0.359	1.112	1.361	0.622
RF	-1.193	-0.237	0.700	-0.924	-0.229	-1.531	-0.094
TD	-0.695	-0.211	-1.057	0.108	0.727	0.874	0.647
LC	-0.245	0.506	-1.505	-0.568	0.487	-0.223	1.027
DP	-0.502	2.082	0.259	-0.475	-0.599	0.532	0.009
	-0.566	**0.523**	**-0.395**	**-0.300**	**0.300**	**0.202**	**0.443**
	0.405	**0.942**	**0.908**	**0.523**	**0.701**	**1.128**	**0.472**
Left Temporal							
CA	-0.392	0.244	-0.922	-0.103	-0.506	0.299	-2.317
SH	0.388	-0.034	-1.393	-1.290	0.326	1.021	-1.980
GM	-0.272	-0.335	-0.801	0.067	0.277	-0.246	-0.441
KW	-0.270	0.079	0.184	-0.601	1.056	1.333	0.685
	-0.137	**-0.011**	**-0.733**	**-0.482**	**0.288**	**0.601**	**-1.014**
	0.354	**0.244**	**0.663**	**0.609**	**0.638**	**0.712**	**1.396**
Right Frontal							

NH	1.612	-0.098	0.477	0.269	-0.285	-0.111	0.538
DF	0.259	-0.321	0.562	-0.134	0.251	0.231	
BS	-0.119	-0.558	0.249	0.340	-0.782	-0.535	0.542
LW	-1.164	-1.759	2.695	-0.473	-1.621	-1.769	0.098
GP	1.830	2.402	2.731	1.643	-1.416	0.622	-2.781
RS	-0.602	-0.122	0.358	0.518	0.510	-0.255	0.594
SN	-1.483	1.089	0.221	1.158	0.161	0.213	0.230
WR	0.418	-0.073	0.481	-0.672	0.144	0.398	0.623
	0.094	**0.070**	**0.972**	**0.331**	**-0.380**	**-0.151**	**-0.022**
	1.198	**1.222**	**1.081**	**0.785**	**0.805**	**0.752**	**1.233**
Left Parietal							
AVG	-0.561	-0.108	-0.822	-0.385	0.739	0.376	0.156
BD	-0.390	-0.797	-0.019	-0.489	0.456	-2.078	1.144
CR	-0.154	-0.057	0.176	-0.245	0.377	-1.988	0.848
NOH	0.838	-0.461	-0.666	1.249	-0.927	-1.305	-0.950
KG	0.349	-0.307	0.107	0.156	0.628	0.735	0.546
SO	1.594	-2.952	-1.236	1.480	-2.755		
	0.279	**-0.780**	**-0.410**	**0.294**	**-0.247**	**-0.852**	**0.349**
	0.823	**1.097**	**0.580**	**0.860**	**1.369**	**1.325**	**0.813**
Right Posterior							
SG	0.077	0.050	-0.370	-0.785	-0.041	-0.107	0.360
AEK	-0.211	-0.230	-0.298	-0.387	0.297	-0.604	0.508
JB	0.006	-0.021	-0.599	-0.309	0.745	0.833	0.195
PC	-0.101	0.607	0.625	-0.233	0.489	0.865	0.154
	-0.057	**0.101**	**-0.160**	**-0.428**	**0.373**	**0.247**	**0.304**
	0.126	**0.357**	**0.539**	**0.246**	**0.331**	**0.725**	**0.163**
Undifferentiated							

	AE	1.180	1.147	0.235	0.728	0.378	1.060	0.185

	Inhibitory Control Measures				Performance Monitoring Measures		
	Antisaccade Accuracy	**Filter proportion with basic**	**Proportion a vs other**	**Stroop Latency difference**	**Errors Design fluency**	**Errors Verbal**	**Post-error Slowing**
Left Frontal							
BP	0.792	-0.021	-0.150	-1.136	0.552	0.862	-1.130
RF	0.792	1.371	-0.228	0.219	1.129	0.673	0.271
TD	-0.076	1.098	-0.807	-0.978	0.874	0.082	
LC	0.358	0.824	-0.171	-0.867	-0.341	0.839	-0.507
DP	0.792	-	-0.686	-0.799	0.988	0.786	-0.677
	0.531	**0.818**	**-0.408**	**-0.712**	**0.640**	**0.648**	**-0.511**
	0.388	**0.602**	**0.313**	**0.536**	**0.588**	**0.325**	**0.584**
Left Temporal							
CA	-0.728	0.277	-0.533	-0.291	0.753	-2.085	-0.533
SH	-1.379	0.642	-0.901	0.921	0.076	-1.191	-0.377
GM	-0.185	0.642	0.589	-0.573	0.351	-1.002	-0.573
KW	0.792	0.025	-0.008	-0.467	-1.322	-1.150	-0.254
	-0.375	**0.397**	**-0.213**	**-0.102**	**-0.036**	**-1.357**	**-0.434**
	0.918	**0.301**	**0.648**	**0.692**	**0.901**	**0.492**	**0.147**
Right Frontal							
NH	0.683	-0.707	-0.768	0.771	-0.054	-0.270	2.227
DF	0.792	-0.191	-0.689	-	1.129	0.839	
BS	0.792	-0.133	0.200	-0.160	-0.929	0.427	0.316

LW	-1.162	-0.379	0.343	-0.479	-1.226	0.345	0.768
GP	-1.270	-2.347	3.055	0.916	-3.284	-0.922	-0.988
RS	0.792	-0.816	-0.420	-0.189	-0.984	-0.783	-0.091
SN	-1.053	-1.568	-1.346	0.714	1.129	0.956	0.859
WR	0.683		-0.565	0.311	-0.984	-0.783	-0.438
	0.032	**-0.877**	**-0.024**	**0.269**	**-0.650**	**-0.024**	**0.379**
	0.991	**0.811**	**1.355**	**0.551**	**1.427**	**0.762**	**1.046**
Left Parietal							
AVG	0.792	-0.191	0.787	-1.166	0.465	-0.439	0.722
BD		-0.953	-0.869	2.091	-0.139	0.896	-0.187
CR	-1.162	0.277	-0.381	0.377	-0.788	0.761	0.011
NOH	-0.945	0.277	0.366	-0.295	-0.444	1.173	0.047
KG	0.358	-0.227	-0.280	-0.193	0.446	1.330	-
SO		-0.379	0.015	2.375	-0.555	0.184	-
	-0.239	**-0.199**	**-0.060**	**0.531**	**-0.169**	**0.651**	**0.148**
	0.960	**0.459**	**0.585**	**1.410**	**0.527**	**0.665**	**0.396**
Right Posterior							
SG	-3.007	-2.066	0.951	-0.764	1.129	0.787	-0.136
AEK	-0.185	-0.816	0.441	-0.035	0.497	-0.500	1.495
JB	0.792	0.642	0.369	-1.607	0.888	-2.619	-0.356
PC	0.792	0.824	0.964	-0.542	-0.524	0.233	-0.132
	-0.402	**-0.354**	**0.681**	**-0.737**	**0.498**	**-0.525**	**0.218**
	1.796	**1.357**	**0.320**	**0.655**	**0.729**	**1.492**	**0.858**
AE	0.683	1.918	0.722	1.845	-0.288	0.324	-0.212

<div align="center">**Supplementary Cognitive Control Analyses:**</div>

S.1. Validating individual measures of cognitive control

As discussed in Chapter 4, our analyses of cognitive control were based on the assumption that the specific measures index a common underlying skill. Here, we perform some exploratory correlational analyses of the relationship between each of the individual measures, which would need to be confirmed on a larger sample.

Activation: Table S.1. presents the Pearson's r values for the individual measures that contribute to the core process of activation.

Table. S1.

Pearson's correlation matrix for the individual measures that contribute to the aggregated measure of activation.

		Reaction-time Speed	Antisaccade RT	Congruent RT
Reaction-time Speed	Correlation	-	0.419*	0.508*
	P Value		*0.03*	*.008*
Antisaccade RT	Correlation		-	-0.185
	P Value			*0.366*

* *p<0.05*

A significant positive correlation was observed between *reaction-time speed* (measure a, Table 4.3) and the following: *Antisaccade RT* (measure c, Table 4.3) and *Congruent RT* (measure b, Table 4.3). Such findings suggest that these measures may indeed be indexing a similar latency construct, despite the fundamentally distinct task requirements.

However, *Antisaccade RT,* and *Congruent RT,* did not significantly correlate with each other. As discussed in Chapter 5, these latency measures involve distinctly different stimulus (and response) modalities, suggesting that such factors may also influence overall latency.

Attentional maintenance: Table S.2. presents the Pearson's r values for the individual measures that contribute to the core process of attentional maintenance.

Table S.2.

Pearson's correlation matrix for the individual measures that contribute to the aggregated measure of attentional maintenance.

		ISI	Prepare	SRT-CRT	Slope
ISI	Correlation	-	0.016	0.036	.459[*]
	P Value		*0.936*	*0.855*	*0.014*
Prepare	Correlation		-	0.226	0.055
	P Value			*0.247*	*0.779*
SRT-CRT	Correlation			-	0.166
	P Value				*0.397*

* *p<0.05*

A significant positive correlation was observed between the *ISI* (inter-stimulus intervals: measure e, Table 4.3) and *slope* (measure d, Table 4.3) measures, suggesting that the two may indeed index a similar, underlying construct. The significant relationship between our slope measure and ISI's appear to support the notion that ISI's may be a better indicator of attentional maintenance.

No significant correlations were observed between *ISI* and *Prepare* (Average warning-time effects: measure g, Table 4.3). This was surprising given both are based on similar constructs of varying intervals between trials. However, this non-significant effect may reflect the restricted intervals used in the warning trial measure: specifically, a long warning tone of three seconds for the *Prepare* is identical in length to a short inter-stimulus interval on the *ISI* measure.

The non-significant effects obtained in the current study are in contrast to observations by Stuss and colleagues (2005), who reported that both warning signals (our warning signal) and task complexity (our SRT-CRT measure) were sensitive to the effects of energisation, particularly in patients with superior medial lesions. This discrepancy may reflect artefacts within the task itself. For example, similar to Stuss and colleagues (2005), warning signals were spaced at one and three seconds prior to the arrival of a stimulus. However, many of our patients exhibited extremely slow response times on the Prepare RT task (e.g. patient SO: 2003ms), thus it is likely that the warning signals were simply too fast for some patients to benefit from. Accordingly, such patients are likely to require higher

thresholds before phasic alertness can be maintained. Upon closer examination, this artefact did not present a difficulty in Stuss and colleagues (2005) study as their patients displayed considerably faster response latencies compared to the current sample.

The above explanation however does not explain the absence of correlation between warning signals and our ISI and slope measures. It may be the case that warning signals better reflect a monitoring component due to the increased expectancy of the upcoming stimulus, which in turn enhances preparation. As such, it may be likely that warning signals themselves act as an external stimulus from which one can modulate their expectancy based. To further investigate this, future studies should compare the relationship between warning effects and other theoretically-motivated measures of monitoring.

Performance Monitoring: Table S.3. presents the Pearson's correlation matrix for our individual measures of attentional maintenance.

Table S.3.

Pearson's correlation matrix for the individual measures that contribute to the aggregated measure of performance monitoring. (PES = Post-error slowing)

		Errors Design Fluency	Errors Verbal Fluency	PES
Errors Design Fluency	Correlation	-	0.154	0.166
	P Value		*0.435*	*0.439*
Errors Verbal Fluency	Correlation		-	0.121
	P Value			*0.574*

* *p<0.05*

There were no significant correlations between any individual measures of performance monitoring. These findings suggest that they may not be indexing a single common construct of performance monitoring. Post-error slowing is theorised to reflect one's ability to notice a deviation from the current goal, and subsequently re-regulate and adjust upcoming behaviour in order to achieve that goal; thus it could be argued that the key component of PES is related *upcoming* behaviours. In contrast, perseverative and inappropriate errors occur due to a failure to monitor the occurrence of *current* responses; therefore, the key component is related to *current* behaviour. Accordingly, it is possible that these are functionally distinct processes (see also Picton et al., 2007).

References

Ahrens, K. (2003). Verbal integration: The interaction of participant roles and sentential arguments. *Journal of Psycholinguistic Research, 32*(5), 497-516. doi:10.1023/A:1025452814385

Alario, F. X., & Cohen, L. (2004). Closed-class words in sentence production: Evidence from a modality-specific dissociation. *Cognitive Neuropsychology, 21*(8), 787-819. doi: 10.1080/02643290342000410

Alexander, M. P. (2006). Impairments of procedures for implementing complex language are due to disruption of frontal attention processes. *Journal of the International Neuropsychological Society, 12*, 236–247. doi: 10.10170S1355617706060309

Alexander, M. P., & Schmitt, M. A. (1980). The aphasia syndrome of stroke in the left anterior cerebral artery territory. *Archives of Neurology, 37*(2), 97-100. doi: 10.1001/archneur.1980.00500510055010

Alexander, M. P., Stuss, D. T., Shallice, T., Picton, T. W., & Gillingham, S. (2005). Impaired concentration due to frontal lobe damage from two distinct lesion sites. *Neurology*, *65*(4), 572-579. doi: 10.1212/01.wnl.0000172912.07640.92

Alexander, M. P., Stuss, D. T., Picton, T., Shallice, T., & Gillingham, S. (2007). Regional frontal injuries cause distinct impairments in cognitive control. *Neurology*, *68*(18), 1515-1523. Doi: 10.1212/01.wnl.0000261482.99569.fb

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264. doi: 10.1016/S0010-0277(99)00059-1

Amieva, H., Phillips, L., & Della Salla, S. (2003). Behavioral dysexecutive symptoms in normal aging. *Brain and Cognition*, *53*(2), 129–132. 10.1016/S0278-2626(03)00094-0

Anderson, S. W., Damasio, H., & Tranel, D. (1990). Neuropsychological impairments associated with lesions caused by tumour or stroke. *Archives of Neurology*, *47*(4), 397– 405. doi:10.1001/archneur.1990.00530040039017

Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, *50*(8), 1787–1793. doi:10.1016/j.neuropsychologia.2012.04.003

Andres, P., & Van der Linden, M. (2001). Supervisory attentional system in patients with focal frontal lesions. *Journal of Clinical and Experimental Neuropsychology, 23*, 225–239. doi: 10.1076/jcen.23.2.225.1212

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition and serial recall: Equivalent measures of verbal short-term memory? *Applied PsychoLinguistics*, *28*(04), 587–606. doi:10.1017/s0142716407070324

Armstrong, C. L., Goldstein, B., Shera, D., Ledakis, G. E., & Tallent, E. M (2003). The predictive value of longitudinal neuropsychological assessment in the early detection of brain tumor recurrence. *Cancer 97*, 649–656. doi:10.1002/cncr.11099

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. *Nature neuroscience*, *6*(2), 115-116. doi: 10.1038/nn1003

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 5-28. doi: 10.1080/713755608

Baddeley, A. (1998). The central executive: A concept and some misconceptions. *Journal of the International Neuropsychological Society*, *4*(5), 523-526. Retrieved from: https://www.cambridge.org

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OUP Oxford.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly journal of experimental psychology*, *36*(2), 233-252. doi:10.1080/14640748408402157

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in cognitive sciences, 12*(5), 193-200. doi: 10.1016/j.tics.2008.02.004

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Cognitive Neuroscience, Journal of*, *19*(12), 2082-2099. doi: 10.1162/jocn.2007.91201

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical?. *Nature Reviews Neuroscience*, *10*(9), 659-669. doi: 10.1038/nrn2667

Badre, D., Hoffman, J., Cooney, J. W., & D'esposito, M. (2009). Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature neuroscience*, *12*(4), 515-522. doi: 10.1038/nn.2277

Bailey, R., & Lucas, T. H. (2014). Impact of Extent of Resection for gliomas. Austin Journal of Surgery, 1(7), 1035. Retrieved from: http://austinpublishinggroup.com/surgery/fulltext/ajs-v1-id1035.php

Baldo, J. V., & Shimamura, A. P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*(2), 259–267. doi:10.1037/0894-4105.12.2.259

Baldo, J. V., Katseff, S., & Dronkers, N. F. (2012). Brain regions underlying repetition and auditory-verbal short-term memory deficits in aphasia: Evidence from voxel-based lesion symptom mapping. *Aphasiology*, *26*(3-4), 338–354. doi: 10.1080/02687038.2011.602391

Baldo, J. V., Schwartz, S., Wilkins, D., & Dronkers, N. F. (2006). Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society, 12*(6), 896–900. doi:10.1017/s1355617706061078

Baldo, J. V., Arévalo, A., Patterson, J. P., & Dronkers, N. F. (2013). Grey and white matter correlates of picture naming: evidence from a voxel-based lesion analysis of the Boston Naming Test. *Cortex*, *49*(3), 658-667. doi:10.1016/j.cortex.2012.03.001

Baldo, J. V., Shimamura, A. P., Delis, D. C., Kramer, J., & Kaplan, E. (2001). Verbal and
design fluency in patients with frontal lobe lesions. *Journal of the International
Neuropsychological Society*, *7*(5), 586-596. Retrieved from: http://ist-
socrates.berkeley.edu/~shimlab/2001_Baldo_Fluency-JINS.pdf

Baldo, J. V., Wilson, S. M., & Dronkers, N. F. (2012). *Uncovering the neural substrates of
language: A voxel-based lesion symptom mapping approach.* Advances in the
Neural Substrates of Language: Toward a Synthesis of Basic Science and Clinical
Research. Oxford: Wiley-Blackwell.

Barbas, H. (1995). Anatomic basis of cognitive-emotional interactions in the primate
prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, *19*(3), 499-510. doi:
10.1016/0149-7634(94)00053-4

Barbey, A. K., Colom, R., & Grafman, J. (2013). Architecture of cognitive flexibility
revealed by lesion mapping. *Neuroimage*, *82*, 547-554. doi:
10.1016/j.neuroimage.2013.05.087

Bartha, L., Knosp, E., Pfisterer, W., & Benke, T. (2000). Intra-and perioperative
monitoring of language functions in patients with tumours in the left perisylvian
area. *Aphasiology, 14*(8), 779-793. doi: 10.1080/026870300412205

Bastiaanse, R., & Jonkers, R. (2012). Linguistic accounts of agrammatic aphasia. In R.
Bastiaansen & C. Thompson (Eds.), *Perspectives on agrammatism* (pp. 17-33).
London: Psychology Press.

Bate, A. J., Mathias, J. L., & Crawford, J. R. (2001). Performance on the Test of Everyday
Attention and standard tests of attention following severe traumatic brain injury.
The Clinical Neuropsychologist, *15*(3), 405-422. doi: 10.1076/clin.15.3.405.10279

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. The
crosslinguistic study of sentence processing, 3, 73-112. Retrieved from:
https://s3.amazonaws.com

Bello, L., Gallucci, M., Fava, M., Carrabba, G., Giussani, C., Acerbi, F., ... & Gaini, S. M.
(2007). Intraoperative subcortical language tract mapping guides surgical removal

of gliomas involving speech areas. *Neurosurgery, 60*(1), 67–80. doi:10.1227/00006123- 200608000-00140

Benton, A, L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia, 6*(1), 53–60. doi: 10.1016/0028-3932(68)90038-9

Benton, A. L. (1969). Development of a multilingual aphasia battery: Progress and problems. *Journal of the neurological sciences, 9*(1), 39-48. doi:10.1016/0022-510X(69)90057-4

Benton, A, L., Hamsher, K, D, S (1983). *Multilingual Aphasia Examination*. Iowa City, IA: AJA Associates

Benton, A. L., Hamsher, K. D. S., Varney, N. R., & Spreen, O. (1983). *Contributions to Neuropsychological Assessment.* New York: Oxford University Press

Berndt, R. S., Burton, M. W., Haendiges, A. N., & Mitchum, C. C. (2002). Production of nouns and verbs in aphasia: effects of elicitation context. *Aphasiology, 16*, 83–106. doi: 10.1080/02687040143000212

Berndt, R. S. (2000) *Quantitative production analysis: A training manual for the analysis of aphasic sentence production*. Psychology Press, 2000.

Berndt, R. S. (2001). More than just words: Sentence production in aphasia. In R. S. Berndt (Ed.), *Handbook of Neuropsychology* (2 ed., Vol. 3, 173-187). Amsterdam: Elsevier.

Berndt, R. S., Burton, M. W., Haendiges, A. N., & Mitchum, C. C. (2002). Production of nouns and verbs in aphasia: Effects of elicitation context. *Aphasiology, 16*(1-2), 83-106. doi: 10.1080/02687040143000212

Berndt, R. S., Haendiges, A. N., Mitchum, C. M., & Sandson, J. (1997). Verb retrieval in aphasia. Relationship to sentence processing. *Brain and Language, 56*(1), 107–137. doi: 10.1006/brln.1997.1728

Berthier, M. L. (2005). Post-stroke aphasia. *Drugs & Aging, 22*(2), 163-182. doi: 10.2165/00002512-200522020-00006

Berman, K. F., & Weinberger, D. R. (1990). Lateralisation of cortical function during cognitive tasks: regional cerebral blood flow studies of normal individuals and patients with schizophrenia. *Journal of Neurology, Neurosurgery & Psychiatry*, *53*(2), 150-160. doi: 10.1136/jnnp.53.2.150

Biassou, N., Obler, L. K., Nespoulous, J. L., Dordain, M., & Harris, K. S. (1997). Dual processing of open-and closed-class words. *Brain and Language, 57*(3), 360-373. doi. 10.1006/brln.1997.1749

Biegler, K. A., Crowther, J. E., & Martin, R. C. (2008). Consequences of an inhibition deficit for word production and comprehension: evidence from the semantic blocking paradigm. *Cognitive Neuropsychology*, *25*(4), 493–527. doi:10.1080/02643290701862316

Bishop, D. V. M. (2003). *The test for Reception for Grammar, version 2 (TROG-2).* London: Psychological Corporation.

Bishop, D.V (1989). *Test for the reception of grammar*: *TROG.* Abingdon, UK: Thomas Leach Ltd.

Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zoltick, B., ... & Mattay, V. S. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *European Journal of Neuroscience*, *23*(6), 1658-1664. doi: 10.1111/j.1460-9568.2006.04680.x

Bock, J. K., & Levelt, W. J. M. (1994). Language production. Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp 945-984). San Diego, CA.

Bogod, N. M., Mateer, C. A., & MacDonald, S. W. (2003). Self-awareness after traumatic brain injury: A comparison of measures and their relationship to executive functions. *Journal of the International Neuropsychological Society, 9*, 450-458. Doi: 10.1017/S1355617703930104

Bogte, H., Flamma, B., van der Meere, J., & van Engeland, H. (2007). Post-error adaptation in adults with high functioning autism. *Neuropsychologia*, *45*(8), 1707-1714. doi: 10.1016/j.neuropsychologia.2006.12.020

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language, 29*(4), 413-432. doi: 10.1016/0749-596X(90)90064-7

Bosma, I., Douw, L., Bartolomei, F., … & Klein, M. (2008). Synchronized brain activity and neurocognitive function in patients with low-grade glioma: A magneto-encephalography study. *Neuro-Oncology, 10*(5),734-744. doi: 10.1215/15228517-2008-034

Bosma, I., Reijneveld, J. C., Klein, M., Douw, L., van Dijk, B. W., Heimans, J. J., & Stam, C. J. (2009). Disturbed functional brain networks and neurocognitive function in low-grade glioma patients: A graph theoretical analysis of resting-state MEG. *Nonlinear Biomedical Physics, 3*(9). doi: 10.1186/1753-4631-3-9

Bosma, I., Vos, M. J., Heimans, J. J., Taphoorn, M. J., Aaronson, N. K., Postma, T. J., van der Ploeg, H. M., Muller, M., Vandertop, W. P., & Slotman, B. J. (2007) The course of neurocognitive functioning in high-grade glioma patients. *Journal of Neuro-Oncology, 9*(1), 53–62. doi: 10.1215/15228517-2006-012

Bosman, F., Carneiro, F., Hruban, R., & Theise, N. (2010). *WHO Classification of Tumours of the Digestive System.* (F. T. Bosman, F. Carneiro, R. H. Hruban, & N. D. Theise, Eds.) *World Health Organization classification of tumours* (p. 417 p.). World Health Organization. Retrieved from http://www.cabdirect.org/abstracts/20113051318.html

Boston Diagnostic Aphasia Examination Booklet: Les and Febiger, 600 South Washington Square, Philadelphia, PA 19106, USA

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 356-366. doi: 10.3758/CABN.7.4.356

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*(5), 201-208. doi: 10.1016/j.tics.2008.02.009

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, *108*(3), 624. doi: 10.1037/0033-295X.108.3.624

Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition*, *113*(3), 262-280. doi: 10.1016/j.cognition.2008.08.011

Bouquet, C. A., Bonnaud, V., & Gil, R. (2003). Investigation of supervisory attentional system functions in patients with Parkinson's disease using the Hayling task. *Journal of Clinical and Experimental Neuropsychology*, *25*(6), 751-760. doi: 10.1076/jcen.25.6.751.16478

Brain Cancer: Gliobastoma multiforme (GBM)." (2013) Retrieved August 2, 2014, from http://www.cancernz.co.nz/brain-cancer

Breckel, T. P., Giessing, C., & Thiel, C. M. (2011). Impact of brain networks involved in vigilance on processing irrelevant visual motion. *NeuroImage*, *55*(4), 1754-1762. doi: 10.1016/j.neuroimage.2011.01.025

Breese, E. L., & Hillis, A. E. (2004). Auditory comprehension: Is multiple choice really good enough? *Brain and Language*, *89*(1), 3-8. doi:10.1016/s0093-934x(03)00412-7

Broca, P. (1861). Nouvelle observation d'aphémie produite par une lésion de la moitié posteriérieure des deuxième et troisième circonvolutions frontales. *Bulletin de la Société Anatomique, 6*, 398–407.

Brown, P. D., Jensen, A. W., Felten, S. J., Ballman, ., V., Schaefer, P. L., Jaeckle, K. A., Cerhan, J. H., & Buckner, J. C. (2006) Detrimental effects of tumor progression on cognitive function of patients with high-grade glioma. *Journal of Clinical Oncology, 24*(34), 5427–5433. doi: 0.1200/JCO.2006.08.5605

Bruyer, R., & Tuyumbu, B. (1980). Verbal fluency and lesion of the cerebral cortex (author's transl). *L'Encephale*, *6*(3), 287-297. Retrieved from: http://europepmc.org/abstract/med/7449726

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*(3), 119–128. doi:10.1016/j.bandl.2010.12.001

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34*, 263-273. doi: 10.1016/0028-3932(95)00104-2

Burgess, P. W., & Shallice, T. (1997). The Hayling and Brixton tests. Retrieved from: http://discovery.ucl.ac.uk/5457

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*(6), 547 – 558. Retrieved from: https://www.cambridge.org/core/journals

Butler, R. W., Rorsman, I., Hill, J. M., & Tuma, R. (1993). The effects of frontal brain impairment on fluency: Simple and complex paradigms. *Neuropsychology*, *7*(4), 519. doi: 10.1037/0894-4105.7.4.519

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108-135). Cambridge, MA: The MIT Press.

Butterworth, B. (1992). Disorders of phonological encoding. *Cognition, 42*(13), 261-286. doi: 10.1016/0010-0277(92)90045-J

Cameron-Jones, C. M. L. (2008). *Lexical Competition Effects in Aphasia: A Thesis Submitted to the Victoria University of Wellington in Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Psychology* (Doctoral dissertation, Victoria University of Wellington).

Canter, G. J., Trost, J. E., & Burns, M. S. (1985). Contrasting speech patterns in apraxia of speech and phonemic paraphasia. *Brain and Language*, *24*(2), 204–222. doi:10.1016/0093-934X(85)90131-2

Caplan, D., & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioural and Brain Sciences*, *22*(1), 77–94. doi:10.1017/s0140525x99001788

Caplan, D., Gow, D., & Makris, N. (1995). Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology*, *45*(2), 293-29. doi:10.1212/wnl.45.2.293

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. doi:10.1080/026432997381664

Caramazza, A., Basili, A. G., Koller, J. J., & Berndt, R. S. (1981). An investigation of repetition and language processing in a case of conduction aphasia. *Brain and language*, *14*(2), 235-271. doi:10.1016/0093-934X(81)90078-X

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and individual differences, 19*(2), 246-251. doi: 10.1016/j.lindif.2008.10.002

Caspari, I., Parkinson, S. R., LaPointe, L. L., & Katz, R. C. (1998). Working memory and aphasia. *Brain and cognition, 37*(2), 205-223. doi: 10.1006/brcg.1997.0970

Chan, R. (2001). Dysexecutive symptoms among a non-clinical sample: A study with the use of the Dysexecutive Questionnaire. *British Journal of Psychology*, *92*(3), 551–565. doi: 10.1348/000712601162338

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201-216. doi: 10.1016/j.acn.2007.08.010

Chapados, C., & Petrides, M. (2013). Impairment only on the fluency subtest of the Frontal Assessment Battery after prefrontal lesions. *Brain, 136*(10), 2966-2978. doi: 10.1093/brain/awt228

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology, 21*, 217-227. doi:10.1016/j.acn.2005.12.002

Christoff, K., Keramatian, K., Gordon, A. M., Smith, R., & Mädler, B. (2009). Prefrontal organization of cognitive control according to levels of abstraction. *Brain research*, *1286*, 94-105. doi: 10.1016/j.brainres.2009.05.096

Clifton, C. (1993). Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology, 47*(2), 222-246. doi: 10.1037/h0078817

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257–271. doi: 10.1093/cercor/bhv165

Cohen, L., Dehaene, S. (1998). Competition between past and present: assessment and interpretation of verbal perseverations. *Brain, 121*(9),1641–1659. doi: 10.1093/brain/121.9.1641

Cohen, R. M., Semple, W. E., Gross, M., Holcomb, H. H., Dowling, M. S., & Nordahl, T. E. (1988). Functional localization of sustained attention: Comparison to sensory stimulation in the absence of instruction. *Cognitive and Behavioral Neurology*, *1*(1), 3-20.

Cohen, R. M., Semple, W, E,. Gross, M., King, A, C., Nordahl, T, E. (1992). Metabolic brain pattern of sustained auditory attention. *Experimental Brain Research, 92*(1), 165–172. Doi: 10.1007/BF00230392

Corbett, F., Jefferies, E., Ehsan, S., & Lambon Ralph, M. A. (2009). Different impairments of semantic cognition in semantic dementia and semantic aphasia: evidence from the non-verbal domain. *Brain, 132*(9), 2593-2608. doi: 10.1093/brain/awp146

Corn, B.W., Wang, M., Fox, S., Michalski, J., Purdy, J., Simpson, J., Kresl, J., Curran, W. J., Diaz, A., Mehta, M., &, Movsas, B. (2009) Health related quality of life and cognitive status in patients with glioblastoma multiforme receiving escalating doses of conformal three dimensional radiation on RTOG 98-03. *Journal of Neuro-Oncology, 95*(2) 247–257. doi: 10.1007/s11060-009-9923-3

Costello, A. B., & Osborne, J. W. (2003). Exploring best practices in factor analysis: Four mistakes applied researchers make. In *Trabajo presentado en la Annual Meeting of the American Educational Research Association (AERA),* Chicago, IL.

Coull, J. T. (1998). Neural correlates of attention and arousal: insights from electrophysiological, functional neuroimaging and psychopharmacology, *Progress in Neurobiology, 55*(4) 343–361. doi: 10.1016/S0301-0082(98)00011-2

Coull, J. T., Frackowiak, R. S. J., & Frith, C. D. (1998). Monitoring for target objects: activation of right frontal and parietal cortices with increasing time on task. *Neuropsychologia*, *36*(12), 1325-1334. doi: 10.1016/S0028-3932(98)00035-9

Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1998). A fronto-parietal network for rapid visual information processing: a PET study of sustained attention and working memory, *Neuropsychologia, 34*(11), 1085–1095. doi: 10.1016/0028-3932(96)00029-2

Crawford, J. R., & Garthwaite, P. H. (2006). Methods of testing for a deficit in single-case studies: Evaluation of statistical power by Monte Carlo simulation. *Cognitive Neuropsychology*, *23*(6), 877-904. doi: 10.1080/02643290500538372

Crawford, J. R., & Howell, D. C. (1998). Payne and Jones Revisited: Estimating the Abnormality of Test Score Differences Using a Modified Paired Samples t Test. *Journal of Clinical and Experimental Neuropsychology, 20*(6), 898– 905. doi:10.1076/jcen.20.6.898.1112

Crosson, B., Sadek, J. R., Maron, L., Gökçay, D., Mohr, C. M., Auerbach, E. J., Freeman, A. J., et al. (2001). Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation. *Journal of Cognitive Neuroscience, 13*(2), 272–283. doi: 10.1162/089892901564225.

Damasio, H. (1998). Neuroanatomical correlates of the aphasias. In M. T. Sarno (Ed.), *Acquired aphasia* (3rd ed., pp. 43–70). San Diego: Academic Press.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *381*(6585), 810-810. doi: 10.1038/381810b0

Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*(1-2), 179–229. doi:10.1016/j.cognition.2002.07.001

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422-433. doi: 10.3758/BF03214546

Darley, F. L., Aronson, A. E., and Brown, J. R. (1975). *Motor Speech Disorders*. Toronto: W. B. Saunders.

Davie, G. L., Hutcheson, K. A., Barringer, D. A., Weinberg, J. S., & Lewin, J. S. (2009). Aphasia in patients after brain tumour resection. *Aphasiology, 23*(9), 1196–1206. doi:10.1080/02687030802436900.

Davies, D. R., & Parasuraman, R. (1982). *The Psychology of Vigilance*. Academic Press. London

Davis, A. (1993). *A Survey of Adult Aphasia, 2nd ed*. Englewood Cliås: Prentice Hall.

de Lacy Costello, A., & Warrington, E. K. (1989). Dynamic aphasia: the selective impairment of verbal planning. *Cortex, 25*(1), 103-114. doi 10.1016/S0010-9452(89)80010-3

De Renzi, E., & Faglioni P. (1978). Normative data and screening power of a shortened version of the token test, *Cortex, 14*(1). 41-49. doi: 10.1016/S0010-9452(78)80006-9

De Witte, E., & Mariën, P. (2013). The neurolinguistic approach to awake surgery reviewed. *Clinical Neurology and Neurosurgery*, *115*(2), 127-145. doi:10.1016/j.clineuro.2012.09.015

De Witte, E., Satoer, D., Robert, E., Colle, H., Verheyen, S., Visch-Brink, E., & Mariën, P. (2015). The Dutch Linguistic Intraoperative Protocol: A valid linguistic approach to awake brain surgery. *Brain and Language, 140*, 35–48. doi: 10.1016/j.bandl.2014.10.011

Değerlendirmesi, B. K. (2012). Comparative assessment of neuro-cognitive impairments among patients with brain tumor and healthy adults. *Turkish Neurosurgery*, *22*(3), 309-316. doi 10.5137/1019-5149.JTN.5144-11.3

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. Psychological Corporation.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321. doi:10.1037/0033-295x.93.3.283.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287–314. doi: 10.1016/0010-0277(92)90046-

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review, 98*, 604–614. doi: 10.1037/0033-295X.98.4.604

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*(2), 125–145. doi:10.1080/02643290342000320

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological review*, *104*(4), 801. doi:10.1037/0033-295x.104.4.801

Deorah, S., Lynch, C. F., Sibenaller, Z. A., & Ryken, T. C. (2006). Trends in brain cancer incidence and survival in the United States: Surveillance, Epidemiology, and End Results Program, 1973 to 2001. *Neurosurgical Focus, 20*(4). Retrieved from: http://thejns.org/doi/abs/10.3171/foc.2006.20.4.e1

Desmurget, M., Bonnetblanc, F., & Duffau, H. (2007). Contrasting acute and slow-growing lesions: a new door to brain plasticity. *Brain, 130*(4), 898-914. doi: 10.1093/brain/awl300

Dijkstra, M., van Nieuwenhuizen, D., Stalpers, L. J., Wumkes, M., Waagemans, M., Vandertop, W. P., ... & Klein, M. (2009). Late neurocognitive sequelea in patients

with WHO grade I meningioma. *Journal of Neurology, Neurosurgery & Psychiatry, 80*(8), 910-915. doi: 10.1136/jnnp.2007.138925

Donati, F., Gobbi, G., Campistol, J., Rapatz, G., Daehler, M., Sturm, Y., et al. (2007) The cognitive effects of oxcarbazepine versus carbamazepine or valproate in newly diagnosed children with partial seizures. *Seizure 16*(8), 670–679. doi: 10.1016/j.seizure.2007.05.006

Dorze, G. L., & Brassard, C. (1995). A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. *Aphasiology, 9*(3), 239-255. doi: 10.1016/j.seizure.2007.05.006

Douw, L., Klein, M., Fagel, S. S., van den Heuvel, J., Taphoorn, M. J., Aaronson, N. K., … & Beute, G. N. (2009). Cognitive and radiological effects of radiotherapy in patients with low-grade glioma: long-term follow up. *The Lancet Neurology, 8*(9) 810-818. doi: 10.1016/S1474-4422(09)70204-2

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature, 384*(6605), 159–161. doi:10.1038/384159a0

Duffau, H., Capelle, L., Denvil, D., Sichez, N., Gatignol, P., Lopes, M., Mitchell, M., Sichez, J., & Van Effenterre, R. (2003). Functional recovery after surgical resection of low-grade gliomas in eloquent brain: hypothesis of brain compensation. *Journal of Neurology, Neurosurgery & Psychiatry, 74*(7), 901–07. doi: 10.1136/jnnp.74.7.901

Duffau, H., Denvil, D., & Capelle, L. (2002). Long term reshaping of language, sensory, and motor maps after glioma resection: a new parameter to integrate in the surgical strategy. *Journal of Neurology, Neurosurgery & Psychiatry*, *72*(4), 511-516. doi: 10.1136/jnnp.72.4.511

Duffau, H. (2005). Lessons from brain mapping in surgery for low-grade glioma: insights into associations between tumour and brain plasticity. *The Lancet Neurology, 4*(8), 476-486. doi: 10.1016/S1474-4422(05)70140-X

Duffau, H. (2007). Contribution of cortical and subcortical electrostimulation in brain glioma surgery: methodological and functional considerations. *Neurophysiologie*

Clinique/Clinical Neurophysiology, 37(6), 373-382.
doi:10.1016/j.neucli.2007.09.003

Duffau, H., Gatignol, P., Mandonnet, E., Capelle, L., & Taillandier, L. (2008).
Intraoperative subcortical stimulation mapping of language pathways in a
consecutive series of 115 patients with Grade II glioma in the left dominant
hemisphere. *Journal of Neurosurgery, 109*(3), 461-471. doi: 10.2217/fnl.10.21

Duffau, H., Gatignol, P., Moritz-Gasser, S., & Mandonnet, E. (2009). Is the left uncinate
fasciculus essential for language? A cerebral stimulation study. *Journal of
Neurology 256*, 382–389. doi: 10.1007/s00415-009-0053-9

Duncan, J. (2006). EPS Mid-Career Award 2004: brain mechanisms of attention. *The
Quarterly Journal of Experimental Psychology*, *59*(1), 2-27. doi:
10.1080/17470210500260674

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental
programs for intelligent behaviour. *Trends in cognitive sciences*, *14*(4), 172-179.
doi: 10.1016/j.tics.2010.01.004

Duncan, J., & Miller, E. K. (2002). Cognitive focus through adaptive neural coding in the
primate prefrontal cortex. *Principles of frontal lobe function*, 278-291. doi:
10.1093/acprof:oso/9780195134971.003.0018

Dutilh, G., van Ravenzwaaij, D., Nieuwenhuis, S., van der Maas, H. L., Forstmann, B. U.,
& Wagenmakers, E. J. (2012). How to measure post-error slowing: a confound and
a simple solution. *Journal of Mathematical Psychology*, *56*(3), 208-216. doi:
10.1016/j.jmp.2012.04.001

Eddy, C. M., Rickards, H. E., & Cavanna, A. E. (2011). The cognitive impact of
antiepileptic drugs. *Therapeutic advances in neurological disorders, 4*(6), 385-407.
doi: 10.1177/1756285611417920

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the
use of exploratory factor analysis in psychological research. *Psychological
methods*, *4*(3), 272. doi: 10.1037/1082-989X.4.3.272

Fadul, C., Wood, J., Thaler, H., Galicich, J,. Patterson, R. H., & Posner,. J. B. (1998). Morbidity and mortality of craniotomy for excision of supratentorial gliomas. *Neurology 38*(9),1374–79. doi: 10.1212/WNL.38.9.1374

Farwell, J. R., Lee, Y. J., Hirtz, D. G., Sulzbacher, S. I., Ellenberg, J. H., & Nelson K. B. (1990). Phenobarbital for febrile seizures: effects on intelligence and on seizure recurrence. *New England Journal of Medicine, 322*, 364–369. doi: 0.1056/NEJM199002083220604

Fassbender, C., Murphy, K., Foxe, J. J., Wylie, G. R., Javitt, D. C., Robertson, I. H., & Garavan, H. (2004). A topography of executive functions and their interactions revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, *20*(2), 132-143. doi: 10.1016/j.cogbrainres.2004.02.007

Faulkner, J. W. (2015). *An evaluation of language in brain tumour patients using a new cognitively-motivated testing protocol. A Thesis Submitted to the Victoria University of Wellington in Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Psychology* (Doctoral dissertation, Victoria University of Wellington).

Faulkner, J. W., Wilshire, C. E., Parker, A. J., & Cunningham, K. (2017). An evaluation of language in brain tumour patients using a new cognitively motivated testing protocol. *Neuropsychology*, *31*(6), 648. doi: 10.1037/neu0000374

Fellows, L. K., & Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control?. *Brain*, *128*(4), 788-796. doi: 10.1093/brain/awh405

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*(3), 348-368. doi: 10.1016/0749-596X(86)90006-9

Field, A. (2005). *Discovering statistics with SPSS*. London, Sage.

Finch, E. & Copland D. A. (2014). Language outcomes following neurosurgery for brain tumours: A systematic review. *Neuro Rehabilitation, 34*(3), 499-514. doi: 10.3233/NRE- 141053

Floden, D., & Stuss, D. T. (2006). Inhibitory control is slowed in patients with right superior medial frontal damage. *Journal of Cognitive Neuroscience, 18*(11), 1843-1849. doi: 10.1162/jocn.2006.18.11.1843

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology Human Perception and Performance, 18*, 1030-1030. doi: 10.1037/0096-1523.18.4.1030

Freedman, M., Alexander, M. P., & Naeser, M. A. (1984). Anatomic basis of transcortical motor aphasia. *Neurology, 34*(4), 409-409. doi: 10.1212/WNL.34.4.409

Frey, R. T., Woods, D. L., Knight, R. T., Scabini, D., & Clayworth, C. (1987). Defining functional areas with averaged CT scans. In *Society for Neuroscience Abstracts* (Vol. 13, p. 1266)

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of experimental psychology: General*, *133*(1), doi: 10.1037/0096-3445.133.1.101

Friedmann, N., Biran, M., & Dotan, D. (2013). Lexical retrieval and breakdown in aphasia and developmental language impairment. *The Cambridge handbook of biolinguistics*, 350-374. Cambridge, UK: Cambridge University Press

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences*, *8*(4), 143-145. doi: 10.1016/j.tics.2004.02.004

Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proceedings of the National Academy of Sciences, 96*(14), 8301-8306. doi: 10.1073/pnas.96.14.8301

Garrard, P., & Hodges, J. R. (2000). Semantic dementia: clinical, radiological and pathological perspectives. *Journal of neurology, 247*(6), 409-422. doi: 10.1007/s004150070169

Garrard, P., Perry, R., & Hodges, J. R. (1997). Disorders of semantic memory. *Journal of Neurology, Neurosurgery, and Psychiatry, 62*(5), 431. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC486839/

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation* (pp. 133–177). New York: Academic Press.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 231-256). Amsterdam: North- Holland Linguistic Series, 30.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177-220). London: Academic Press.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213. doi:10.1016/0749-596X(89)90044-2

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*. doi:10.1037/0012-1649.28.5.887

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*, *64*, 310-319. doi: 10.1016/j.neuropsychologia.2014.09.024

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science*, *4*(6), 385-390. doi: 10.1111/j.1467-9280.1993.tb00586.x

Gehring, K., Sitskoorn, M. M., Aaronson, N. K., & Taphoorn, M. J. B. (2008). Interventions for cognitive deficits in adults with brain tumours. *The Lancet Neurology, 7*(6), 548-560. doi: 10.1016/S1474-4422(08)70111-X

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430. doi: 10.1037/0278-7393.16.3.430

Geschwind, N. (1965). Disconnexion syndromes in animals and man, *Brain, 88*(2), 237–294. doi: 10.1093/ brain/88.2.237.

Gillham, R. A., Read, C. L., McKee, P. J. M., Larkin, J. G., & Brodie, M. (1991) Cognitive function in epileptic patients on long-term sodium valproate. *Journal of Epilepsy, 4*(4), 205–210. doi: 10.1016/0896-6974(91)90077-V

Giovagnoli, A. R., Casazza, M., Ciceri, E., & Avanzini, G. (2006) Preserved memory in temporal lobe epilepsy patients after surgery for low-grade tumour. A pilot study. *Neurological Sciences 28*(5):171–174. doi: 10.1007/s10072-007-0831-z

Giovagnoli, A. R. (2012). Investigation of cognitive impairments in people with brain tumors. *Journal of Neuro-Oncology, 108*(2), 277-283. doi: 10.1007/s11060-012-0815-6

Giovagnoli, A. R., Silvani, A., Colombo, E., & Boiardi, A. (2005). Facets and determinants of quality of life in recurrent high-grade glioma patients. *Journal of Neurology, Neurosurgery, & Psychiatry 76*(4), 562–568. doi: 10.1136/jnnp.2004.036186

Godefroy, O., Cabaret, M., Petit-Chenal, V., Pruvo, J. P., & Rousseaux, M. (1999). Control functions of the frontal lobes. Modularity of the central-supervisory system?. *Cortex*, *35*(1), 1-20. doi: 10.1016/S0010-9452(08)70782-2

Goldberg, G., Mayer, N. H., & Toglia, J. U. (1981). Medial frontal cortex infarction and the alien hand sign. *Archives of Neurology, 38*(11), 683-686. doi: 10.1001/archneur.1981.00510110043004

Goldstein, B., Obrzut, J. E., John, C., Ledakis, G., & Armstrong, C. L. (2004). The impact of frontal and non-frontal brain tumor lesions on Wisconsin Card Sorting Test performance. *Brain and Cognition*, *54*(2), 110-116. doi: 10.1016/S0278-2626(03)00269-0

Goldstein, G. (1996). Functional considerations in neuropsychology. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 75–89). Delray Beach, Florida: GR Press/St. Lucie Press.

Goodglass, H., & Kaplan, E. (1983). *Boston diagnostic aphasia examination booklet*. Lea & Febiger.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.

Grant, R., Slattery, J., Gregor, A., & Whittle, I. R (1994). Recording neurological impairment in clinical trials of glioma. *Journal of Neuro-Oncology, 19*(1),37–49. doi: 10.1007/BF01051047

Green, J. B. (2003). Brain reorganisation after stroke. *Topics in Stroke Rehabilitation, 10*(3), 1-20. doi: 10.1310/H65X-23HW-QL1G-KTNQ

Grimshaw, J. (2000). *Argument Structure*. Cambridge, MA: MIT Press

Gulati, S., Berntsen, E. M., Solheim, O., Kvistad, K. A., Håberg, A., Selbekk, T., ... & Unsgaard, G. (2009). Surgical resection of high-grade gliomas in eloquent regions guided by blood oxygenation level dependent functional magnetic resonance imaging, diffusion tensor tractography, and intraoperative navigated 3D ultrasound. *Minimally Invasive Neurosurgery*, *52*(01), 17-24. Retrieved from: https://profile.thieme.de/HTML/sso/ejournals

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *56*(7), 1213–1236. doi:10.1080/02724980343000071

Gupta, P., MacWhinney, B., Feldman, H. M., & Sacco, K. (2003). Phonological memory and vocabulary learning in children with focal lesions. *Brain and Language*, *87*(2). 241– 252. doi:10.1016/S0093-934X(03)00094-4

Gvion, A., & Friedmann, N. (2012). Does phonological working memory impairment affect sentence comprehension? A study of conduction aphasia. *Aphasiology*. doi:10.1080/02687038.2011.647893

Haas, J., Vogt, G., Schiemann, M., & Patzold, U. (1982). Aphasia and non-verbal intelligence in brain tumour patients. *Journal of Neurology, 227*(4), 209-218. doi: 10.1007/BF00313388

Haglund, M. M., Berger, M. S., Shamseldin, M., Lettich, E., & Ojemann, G. A. (1994). Cortical localization of temporal lobe language sites in patients with gliomas. *Neurosurgery*, *34*(4), 567–576. doi:10.1097/00006123-199404000-00001

Hahn, C.A., Dunn, R.H., Logue, P.E., King, J.H., Edwards, C.L., & Halperin, E.C. (2003). Prospective study of neuropsychological testing and quality-of-life assessment of adults with primary malignant brain tumours. *International Journal of Radiation, Oncology, Biology, Physics, 55*(4). 992–999. doi: 10.1016/S0360-3016(02)04205-0

Hamilton, A. C., & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single-case study. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(1), 1-13. doi:10.3758/CABN.5.1.1

Harder, H., Holtel, H., Bromberg, J. E. C., Poortmans, P., Haaxma-Reiche, H., Kluin-Nelemans, H. C., Menten, J., & van den Bent, M. J (2004.) Cognitive status and quality of life for primary CNS lymphoma. *Neurology 62*(4). 544–547. doi: 10.1212/WNL.62.4.544

Harley, T. (2001). *The Psychology of Language. From Data to Theory* (second edition). Psychology Press. East Sussex

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science, 8*(3), 191-219. doi: 10.1207/s15516709cog0803_1

Heimans, J. J., & Reijneveld, J. C. (2012). Factors affecting the cerebral network in brain tumor patients. *Journal of Neuro-Oncology*, *108*(2), 231-237. doi:10.1007/s11060-012-0814-7

Heimans, J. J., & Taphoorn, M. J. (2002). Impact of brain tumour treatment on quality of life. *Journal of neurology*, *249*(8), 955-960. doi: 10.1007/s00415-002-0839-5

Henseler, I., Regenbrecht, F., & Obrig, H. (2014). Lesion correlates of patholinguistic profiles in chronic aphasia: comparisons of syndrome-, modality- and symptom-level assessment. *Brain, 137*(3), 918–930. doi:10.1093/brain/awt374

Hickok, G., & Poeppel, D. (2007). Opinion - The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. doi:10.1038/nrn2113

Hillis, A. E., Kleinman, J. T., Newhart, M., Heidler-Gary, J., Gottesman, R., Barker, P. B., Aldrich, E., Llinas, R., Wityk, R., & Chaudhry, P. (2006). Restoring cerebral blood flow reveals neural regions critical for naming. *Journal of Neuroscience, 26*(31), 8069-8073. doi: 10.1523/JNEUROSCI.2088-06.2006

Hodges, J. R., & Patterson, K. (2007). Semantic dementia: a unique clinicopathological syndrome. *The Lancet Neurology*, *6*(11), 1004-1014. doi:10.1016/s1474-4422(07)70266-1

Hom, J., & Reitan, R. M. (1984). Neuropsychological correlates of rapidly vs. slowly growing intrinsic cerebral neoplasms. *Journal of Clinical and Experimental Neuropsychology, 6*(3), 309-324. doi: 10.1080/01688638408401221

Houx, P. J., Jolles, J. (1994). Vulnerability factors for age-related cognitive decline. In: Isaacson, R. L., & Jensen, K. F. (Eds), *The vulnerable brain and environmental risks* (pp 25 – 41). New York: Plenum Press

Huber, W., Poeck, K., & Willmes, K. (1984). The Aachen Aphasia test. *Advances in Neurology, 42.* 291-303.

Hula, W. D., & McNeil, M. R. (2008). Models of attention and dual-task performance as explanatory constructs in aphasia. In *Seminars in Speech and Language*, *29*(03), 169-187. Thieme Medical Publishers. doi: 10.1055/s-0028-1082882

Illmberger, J., Ruge, M., Kreth, F. W., Briegel, J., Reulen, H. J., & Tonn, J. C. (2008). Intraoperative mapping of language functions: A longitudinal neurolinguistic analysis. *Journal of Neurosurgery, 109*(4), 583-592. Retrieved from: http://thejns.org/doi/full/10.3171/JNS/2008/109/10/0583

Imperato, J. P., Paleologos, N. A., & Vick, N. A. (1990). Effects of treatment on long-term survivors with malignant astrocytomas. *Annals of Neurology, 28*(6), 818-822. doi: 10.1002/ana.410280614

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1-2), 101–144. doi:10.1016/j.cognition.2002.06.001

Ito, M., Hatazawa, J,. Yamaura, H., & Matsuzawa, T. (1981). Age-related brain atrophy and mental deterioration - a study with computed tomography. *British Institute of Radiology, 54*(641), 384-90. doi: 10.1259/0007-1285-54-641-384

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences, 10*, 480-486. doi: 10.1016/j.tics.2006.09.002

Jakola AS, Myrmel KS, Kloster R, Torp SH, Lindal S, Unsgard G, et al. Comparison of a strategy favoring early surgical resection vs a strategy favoring watchful waiting in low-grade gliomas. JAMA. 2012; 308: 1881-1888

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience, 21*(12), 2434–2444. doi: 10.1162/jocn.2008.21179

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain: A Journal of Neurology*, *129*(8), 2132–2147. doi:10.1093/brain/awl153

Johns, D. F., & Darley, F. L. (1970). Phonemic variability in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *13*(3), 556-583. doi:10.1044/jshr.1303.556

Jones-Gotman, M., & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal cortical lesions. *Neuropsychologia*, *15*(4), 653-674. doi: 10.1016/0028-3932(77)90070-7

Jongman, S. R., Roelofs, A., & Meyer, A. S. (2015). Sustained attention in language production: An individual differences investigation. *The Quarterly Journal of Experimental Psychology*, *68*(4), 710-730. doi: 10.1080/17470218.2014.964736

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive science, 20*(2), 137-194. doi: 10.1207/s15516709cog2002_1

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. doi: 10.1007/BF02291575

Kaleita, T. A., Wellisch, D. K., Cloughesy, T. F., Ford, J. M., Freeman, D., Belin, T. R., & Goldman, J. (2004). Prediction of neurocognitive outcome in adult brain tumor patients. *Journal of Neuro-Oncology, 67*(1), 245-253. doi: 10.1023/B:NEON.0000021900.29176.58

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test*. Philadelphia: Lippincott, Williams and Wilkins

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test (2nd ed).* Philadelphia: Lea & Febiger

Karnath, H. O., & Steinbach, J. P. (2011). Do brain tumours allow valid conclusions on the localisation of human brain functions?–Objections. *Cortex*, *47*(8), 1004-1006. doi: 10.1016/j.cortex.2010.08.006

Katzev, M., Tüscher, O., Hennig, J., Weiller, C., Kaller, C. P. (2013). Revisiting the functional specialization of left inferior frontal gyrus in phonological and semantic fluency: the crucial role of task demands and individual ability. *Journal of Neuroscience, 33*(18), 7837–7845 doi: 10.1523/JNEUROSCI.3147-12.2013

Kauhanen, M. L., Korpelainen, J. T., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., ... & Myllylä, V. V. (2000). Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. *Cerebrovascular Diseases Basel Switzerland*, *10*(6), 455–461. doi:10.1159/000016107

Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. *Aphasiology*, *10*(2), 159-180. doi:10.1080/02687039608248403

Keles, G. E., Lundin, D. A., Lamborn, K. R., Chang, E. F., Ojemann, G., & Berger, M. S (2004). Intraoperative subcortical stimulation mapping for hemispherical perirolandic gliomas located within or adjacent to the descending motor pathways: evaluation of morbidity and assessment of functional outcome in 294 patients. *Journal of Neurosurgery, 100*(3), 369–375. Retrieved from: http://thejns.org/doi/full/10.3171/jns.2004.100.3.0369

Kerns, J. G., Cohen, J. D., MacDonald III, A. W., Johnson, M. K., Stenger, V. A., Aizenstein, H., & Carter, C. S. (2005). Decreased conflict-and error-related activity in the anterior cingulate cortex in subjects with schizophrenia. *American Journal of Psychiatry*, *162*(10), 1833-1839. doi: 10.1176/appi.ajp.162.10.1833

Kertesz, A. (1982). *Western Aphasia Battery test manual.* New York: Grune and Stratton.

Khan, F., & Amatya, B. (2013). Factors associated with long-term functional outcomes, psychological sequelae and quality of life in persons after primary brain tumour. *Journal of Neuro-Oncology, 111*(3), 335-366. doi: 10.1007/s11060-012-1024-z

Khng, K. H., & Lee, K. (2009). Inhibiting interference from prior knowledge: Arithmetic intrusions in algebra word problem solving. *Learning and Individual Differences*, *19*(2), 262-268. doi: 10.1016/j.lindif.2009.01.004

Kiernan, R. J., Mueller, J., Langston, J. W., & Van Dyke, C. (1987). The neurobehavioural cognitive status examination: a brief but quantitative approach to cognitive assessment. *Annals of Internal Medicine, 107*(4), 481-485. doi: 10.7326/0003-4819-107-4-481

Klein, M., & Heimans, J. J. (2004). The measurement of cognitive functioning in low-grade glioma patients after radiotherapy. *Journal of Clinical Oncology, 22*(5), 966-967. doi 10.1200/JCO.2004.99.290

Klein, M., Engelberts, N. H., van der Ploeg, H. M., Kasteleijn-Nolst Trenité, D. G., Aaronson, N. K., Taphoorn, M. J., ... & Heimans, J. J. (2003). Epilepsy in low-grade gliomas: The impact on cognitive function and quality of life. *Annals of neurology*, *54*(4), 514-520. doi: 10.1002/ana.10712

Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, *50*(6), 963-974. doi: 10.1016/j.neuron.2006.05.017

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*(6), 229-235. /doi: 10.1016/j.tics.2007.04.005

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*(5648), 1181-1185. doi: 10.1126/science.1088545

Kośla, K., Pfajfer, L., Bryszewski, B., Jaskólski, D., Stefańczyk, L., & Majos, A. (2012). Functional rearrangement of language areas in patients with tumors of the central nervous system using functional magnetic resonance imaging. *Polish Journal of Radiology, 77*(3), 39. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3447432/

Lambon R. M. A., Sage, K., & Roberts, J. (2000). Classical anomia: A neuropsychological perspective on speech production. *Neuropsychologia*, *38*(2), 186- 202. doi:10.1016/s0028-3932(99)00056-1

Larrabee, G. J., Trahan, D. E., Curtiss, G., & Levin, H. S (1988). Normative data for the Verbal Selective Reminding Test. *Neuropsychology,2*(3-4), 173-182. doi: 10.1037/h0091731

Lee, G., Strauss, E., McCloskey, L., Loring, D., & Drane, D. (1996). Localization of frontal lobe lesions using verbal and nonverbal fluency measures. In *Annual Meeting of the International Neuropsychological Society,* Chicago.

Lee, V. E. (2017). A "core skills" approach to the assessment of acquired language disorders: Exploration and cross-validation. *A Thesis Submitted to the Victoria University of Wellington in Fulfilment of the Requirements for the Degree of Masters of Science in Cognitive and Behavioural Neuroscience* (Masters Thesis, Victoria University of Wellington).

Lehto, J. (1996). Are executive function tests dependent on working memory capacity?. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 29-50. doi: 10.1080/713755616

Levelt, W. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. doi:10.1016/S1364-6613(99)01319-4

Levelt, W. J. M. (1989) *Speaking: From intention to articulation*. Bradford, Cambridge, MA: The MIT Press

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75. doi: 10.1017/S0140525X99001776

Levin, H. S., Fletcher, J. M., Kufera, J. A., Harward, H., Lilly, M. A., Mendelsohn, D., ... & Eisenberg, H. M. (1996). Dimensions of cognition measured by the tower of London and other cognitive tasks in head-injured children and adolescents. *Developmental Neuropsychology*, *12*(1), 17-34. doi: 10.1080/87565649609540638

Lewin, J. S., Friedman, L., Wu, D., Miller, D. A., Thompson, L. A., Klein, S. K., ... & Friedland, R. P. (1996). Cortical localization of human sustained attention: detection with functional MR using a visual vigilance paradigm. *Journal of computer assisted tomography*, *20*(5), 695-701.

Lezak, M. D. (2004). Neuropsychological assessment. New York: Oxford University Press

Lichtheim, L. (1885). On aphasia. *Brain, 7*, 433–484. doi: 10.1093/brain/7.4.433

Lieu, A.S., & Howng, S. L. (2000). Intracranial meningiomas and epilepsy: incidence, prognosis and influencing factors. *Epilepsy Research, 38*(1), 45–52. doi: 10.1016/S0920-1211(99)00066-2

Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A., ... & Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica, 114*(2), 97-109. doi: 10.1007/s00401-007-0243-4

Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., ... & Ellison, D. W. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica, 131(*6), 803-820. doi: 10.1007/s00401-016-1545-1

Lowe, C., & Rabbitt, P. (1997). Cognitive models of aging and frontal lobe deficits. In Rabbit, R. *Methodology of frontal and executive function* (pp 39-59). Psychology Press

Luteijn, F., Barelds, D. P. F (2004). *Groninger intelligentie test II* (GIT II). Pearson, Amsterdam

Luo L., Luk G., & Bialystok E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition 114*(1), 29–41 10.1016/j.cognition.2009.08.014

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review, 101*(4), 676. Doi: 10.1080/01690969408402115

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*(5472), 1835-1838. doi: 10.1126/science.288.5472.1835

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*(2), 157-201. doi: 10.1080/01690969408402115

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, *109*(2), 163-203. doi: 10.1037/0033-2909.109.2.163

MacPherson, S. E., Turner, M. S., Bozzali, M., Cipolotti, L., & Shallice, T. (2010). Frontal subregions mediating Elevator Counting task performance. *Neuropsychologia, 48*(12), 3679-3682. doi: 10.1016/j.neuropsychologia.2010.07.033

MacWinney, B. (1991). *The CHILDES project: tools for analyzing talk*. Lawrence Erlbaum Associates Inc, Hillsdale

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology, 33*(3), 503–535. doi: 10.1037/0278-7393.33.3.503

Manni, R., Ratti, M.T., Perruca, E., Galimberti, C.A., Tartara, A. (1993) A multiparametric investigation of daytime sleepiness and psychomotor functions in patients treated with phenobarbital and sodium valproate – a comparative controlled study. *Electroencephalography and Clinical Neurophysiology, 86*(5), 322–328. doi: 10.1016/0013-4694(93)90044-V

Marshall, J., Pring, T., & Chiat, S. (1998). Verb retrieval and sentence production in aphasia. *Brain and Language*, *63*(2), 159-183. doi:10.1006/brln.1998.1949

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1), 71-102. doi:10.1016/0010-0277(87)90005-9

Martin, N., & Gupta, P. (2004). Exploring the relationship between word processing and verbal short term memory: Evidence from association and dissociation. *Cognitive Neuropsychology, 21(2-4)*, 213–228. doi: 10.1080/02643290342000447

Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology, 14*(5), 641-682. doi: 10.1080/026432997381402

Martin, R. C., & Feher, E. (1990). The consequences of reduced memory span for the comprehension of semantic versus syntactic information. *Brain and Language, 38*(1), 1-20. doi: 10.1016/0093-934X(90)90099-3

Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology, 8*(4), 506. doi: 10.1037/0894-4105.8.4.506

Martin, R. C., Loring, D. W., Meador, K. J., & Lee, G. P. (1990). The effects of lateralized temporal lobe dysfunction on normal and semantic word fluency. *Neuropsychologia*, *28*(8), 823-829. doi: 0.1016/0028-3932(90)90006-A

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: past studies and new evidence. *Cortex, 45*(6), 738-758. doi: 10.1016/j.cortex.2008.10.003

Mazaux, J, M., & Orgogozo, J, M. (1982). Echelle d'e ́valuation de l'aphasie adapte ́e du *Boston Diagnostic Aphasia Examination*. E.A.P. Editions Psychotechniques.

McAleer, M. F., & Brown, P. D. (2015). Therapeutic management of gliosarcoma in the temozolomide era. *CNS Oncology, 4*(3), 171-178. doi: 10.2217/cns.14.61

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, *109*(2), 163. doi: 10.1037/0033-2909.109.2.163

McKee, P. J. W., Blacklaw, J., Butler, E., Gillham, R. A., & Brodie, M. J. (1992) Variability and clinical relevance of the interaction between sodium valproate and carbamazepine in epileptic patients. *Epilepsy Research, 11*(3), 193–198. doi: 10.1016/0920-1211(92)90098-E

Meador, K.J., Loring, D. W., Ray, P. G., Murro, A. M., King, D. W., Nichols, M. E., Deer, E. W., & Goff, W. T. (1999). Differential cognitive effects of carbamazepine and gabapentin. *Epilepsia, 40*(9). 1279–1285. doi: 10.1111/j.1528-1157.1999.tb00858

Metz-lutz, M, N., Kremin, H., Deloche, G., Hannequin, D., Ferrand, I., & Perrier, D (1991). Standardisation d'un test de de ́nomination orale: controˆle des effets de l'aˆge, du sexe et du niveau de scolarite ́chez les sujets adultes normaux. *Review of Neuropsychology, 1*(1), 73–95.

Meyers, C. A., & Brown, P. D. (2006). Role and relevance of neurocognitive assessment in clinical trials of patients with CNS tumours. *Journal of Clinical Oncology*, *24*(8), 1305–1309. doi:10.1200/jco.2005.04.6086

Meyers, C. A., & Hess, K. R (2003). Multifaceted end points in brain tumor clinical trials: cognitive deterioration precedes MRI progression. *Neuro-Oncology 5*(2), 89–95. doi: 10.1093/neuonc/5.2.89

Meyers, C. A., Hess, K. R., Yung, W. K., & Levin, V.A. (2000). Cognitive function as a predictor of survival in patients with recurrent malignant glioma. *Journal of Clinical Oncology, 18*(3), 646–50. doi: 10.1200/JCO.2000.18.3.646

Miceli, G., Amitrano, A., Capasso, R., & Caramazza, A. (1996). The treatment of anomia resulting from output lexical damage: Analysis of two cases. *Brain and Language*, *52*(1), 150-174. doi:10.1006/brln.1996.0008

Miceli, G., Caltagirone, C., Gainotti, G., Masullo, C., & Silveri, M. C. (1981). Neuropsychological correlates of localized cerebral lesions in non-aphasic brain-damaged patients. *Journal of Clinical and Experimental Neuropsychology, 3*(1), 53-63. doi: 10.1080/01688638108403113

Miceli, G., Capasso, R., Monti, A., Santini, B., & Talacchi, A. (2012). Language testing in brain tumor patients. *Journal of Neuro-Oncology, 108*(2), 247-252.doi:10.1007/s11060-012-0810-y.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167-202. doi: 10.1146/annurev.neuro.24.1.167

Miller, L. S., & Rohling, M. L. (2001). A statistical interpretive method for neuropsychological test data. *Neuropsychology Review*, *11*(3), 143–169. doi: 10.1023/A:1016602708066

Ministry of Health (2010). *Cancer Projections: Incidence 2004 - 2008 to 2014 - 2018*. Wellington, Ministry of Health.

Miotto, E. C., Silva Junior, A., Silva, C. C., Cabrera, H. N., Machado, M. A., Benute, G. R., ... & Teixeira, M. J. (2011). Cognitive impairments in patients with low grade gliomas and high grade gliomas. *Arquivos de Neuro-Psiquiatria, 69*(4), 596-601. doi: 10.1590/S0004-282X2011000500005

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. *Cognitive psychology*, *41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Monje, M. L., Mizumatsu, S., Fike, J. R., & Palmer, T. D. (2002). Irradiation induces neural precursor-cell dysfunction. Nature Medicine, 8(9), 955–62. doi: 10.1038/nm749

Moritz-Gasser, S., Herbet, G., Maldonado, I. L., & Duffau, H. (2012). Lexical access speed is significantly correlated with the return to professional activities after awake surgery for low-grade gliomas. *Journal of Neuro-Oncology*, *107*(3), 633-641. doi: 10.1007/s11060-011-0789-9

Murray, L. L. (2012). Attention and other cognitive deficits in aphasia: Presence and relation to language and communication measures. *American Journal of Speech-Language Pathology*, *21*(2), S51-S64. doi: 10.1044/1058-0360(2012/11-0067)

Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., & Parasuraman, R. (2014). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *Neuroimage*, *85*, 909-917. doi: 10.1016/j.neuroimage.2012.11.061

Nickels, L. (2002). Theoretical and methodological issues in the cognitive neuropsychology of spoken word production. *Aphasiology*, *16*(1-2), 3-19. doi:10.1080/02687040143000645

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220-246. doi.org/10.1037/0033-2909.126.2.220

Noll, K. R., Sullaway, C., Ziu, M., Weinberg, J. S., & Wefel, J. S. (2014). Relationships between tumor grade and neurocognitive functioning in patients with glioma of the left temporal lobe prior to surgical resection. *Neuro-Oncology, 17*(4), 580–7. doi: 10.1093/neuonc/nou233 84

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18). Springer: USA.

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition*, *111*(2), 275-279. doi: 10.1016/j.cognition.2009.02.002

Novelli, G., Papagno, C., Capitani, E., Laiacona, M., Vallar, G., & Cappa S, F. (1986). Tre test clinici di ricerca e produzione lessicale: taratura su soggetti normali, *Archivio di Neurologia, Psicologia e Psichiatria , 47*(4), 477-506. doi: 10.1017/S1355617711001676

Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology, 26*(6), 527-567. doi: 10.1080/02643290903519367

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 263-281. doi: 10.3758/cabn.5.3.263

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, *4*(10), 906-924. doi: 10.1111/j.1749-818x.2010.00244.x

Nyhus, E. & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. Brain and Cognition, 71, 437-451. doi: 10.1016/j.bandc.2009.03.005

Ogar, J. M., Dronkers, N. F., Brambati, S. M., Miller, B. L., & Gorno-Tempini, M. L. (2007). Progressive nonfluent aphasias and its characteristic motor speech deficits, *Alzheimers Disease and Associated Disorders, 21*(4), doi: 10.1097/WAD.0b013e31815d19fe

Olsen, R., Iverson, G., Carolan, H., Parkinson, M., Brooks, B., McKenzie, M. (2011). Prospective comparison of two cognitive screening tests; diagnostic accuracy and

correlation with community integration and quality of life. *Journal of Neuro-Oncology, 105*, 337–44. doi:10.1007/s11060- 011- 0595- 4

Ostrom, Q, T., Gittleman, H., de Blank, P, M., Finlay, J, L., Gurney, J, G., McKean-Cowdin, R., … & Barnholtz-Sloan, J, S. (2016). American Brain Tumor Association Adolescent and Young Adult Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2008-2012. *Neuro Oncology, 18*(1). doi: https://doi.org/10.1093/neuonc/nov297.

Ostrom, Q, T., Gittleman, H., Liao, P., Vecchione-Koval, T., Wolinsky, Y., Kruchko, C., Barnholtz-Sloan, J, S. (2017). CBTRUS Statistical report: Primary brain and other central nervous system tumours diagnosed in the United States in 2010-2014. *Neuro-Oncology, 19*(5), v1-v88. Doi: 10.1093/neuonc/nox158

Ostrom, Q. T., Bauchet, L., Davis, F. G., Deltour, I., Fisher, J. L., Langer, C. E., ... & Wrensch, M. R. (2014). The epidemiology of glioma in adults: a "state of the science" review. *Neuro-Oncology, 16*(7), 896-913. doi 10.1093/neuonc/nou087

Ostrom, Q. T., Gittleman, H., Fulop, J., Liu, M., Blanda, R., Kromer, C., ... & Barnholtz-Sloan, J. S. (2015). CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008-2012. *Neuro-Oncology, 17*(Suppl 4), iv1.-iv62. doi:10.1093/neuonc/nov189.

Påhlson, A., Ek, L., Ahlström, G., & Smits, A. (2003). Pitfalls in the assessment of disability in individuals with low-grade gliomas. *Journal of Neuro-Oncology, 65*(2), 149–158. doi:10.1023/b:neon.0000003727.09448.dd

Pai, M. C. (1999). Supplementary motor area aphasia: a case report. *Clinical Neurology and Neurosurgery, 101*(1), 29-32. doi: 10.1016/S0303-8467(98)00068-7

Papagno, C., Casarotti, A., Comi, A., Gallucci, M., Riva, M., & Bello, L. (2012). Measuring clinical outcomes in neuro-oncology. A battery to evaluate low-grade gliomas (LGG). *Journal of Neuro-Oncology, 108*(2), 269–75. doi:10.1007/s11060-012-0824-5

Papagno, C. and Basso, A. (1996). Perseveration in two aphasic patients. *Cortex, 32*, 67–
82. doi: 10.1016/S0010-9452(96)80017-7

Papagno, C., Cecchetto, C., Reati, F., & Bello, L. (2007). Processing of syntactically
complex sentences relies on verbal short-term memory: evidence from a short-term
memory patient. *Cognitive Neuropsychology, 24*(3), 292-311. doi:
10.1080/02643290701211928.

Papagno, C., Miracapillo, C., Casarotti, A., Romero Lauro, L. J., Castellano, A., Falini, A.,
... & Bello, L. (2010). What is the role of the uncinate fasciculus? Surgical removal
and proper name retrieval. *Brain, 134*(2), 405-414. doi: 10.1093/brain/awq283

Papaioannou, A., Fraidakis, O., Michaloudis, D., Balalis, C., & Askitopoulou, H. (2005).
The impact of the type of anesthesia on cognitive status and delirium during the
first postoperative days in elderly patients. *European Journal of Anesthesiology,
22*(7), 492-499. doi.org/10.1017/S0265021505000840

Parisi, D., & Pizzamiglio, L. (1970). Syntactic comprehension in aphasia, *Cortex, 6*(2),
204-15. doi: 10.1016/S0010-9452(70)80028-4

Park, B. J, Kim, H. K., Sade, B., & Lee, J. H. (2009). "Epidemiology". In Lee J. H.
Meningiomas: Diagnosis, Treatment and Outcome. *American Journal of
Neuroradiology, 30*(10), doi: 10.3174/ajnr.A1775

Pashek, G. V, & Holland, A. L. (1988). Evolution of aphasia in the first year post-onset.
Cortex, *24*(3), 411–423. doi:10.1016/s0010-9452(88)80004-2

Pate, D. S., Saffran, E. M., & Martin, N. (1987). Specifying the nature of the production
impairment in a conduction aphasic: A case study. *Language and Cognitive
Processes*, *2*(1), 43-84. doi:10.1080/01690968708406351

Pedersen, P. M., Vinter, K., & Olsen, T. S. (2004). Aphasia after stroke: type, severity and
prognosis. *Cerebrovascular Diseases, 17*(1), 35-43. doi: 10.1159/000073896

Pelletier, G., Verhoef, M. J., Khatri, N., & Hagen, N. (2002). Quality of life in brain tumor
patients: the relative contributions of depression, fatigue, emotional distress, and

existential issues. *Journal of Neuro-Oncology*, *57*(1), 41-49. doi: 10.1023/A:1015728825642

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*(4), 633-651. doi: 10.1006/jmla.1998.2592

Picton, T. W, Stuss, D. T., Alexander, M. P., Shallice, T., Binns, M. A., & Gillingham, S. (2007). Effects of frontal lesions on response inhibition. Cerebral Cortex, 17(4), 826-838. doi: 10.1093/cercor/bhk031

Pijnenburg, Y. A. L., Gillissen, F., Jonker, C., & Scheltens, P. (2004). Initial complaints in frontotemporal lobar degeneration. *Dementia and Geriatric Cognitive Disorders*, *17*(4), 302–306. doi:10.1159/000077159

Piras, F., & Marangolo, P. (2007). Noun-verb naming in aphasia: a voxel-based lesion-symptom mapping study. *Neuroreport*, *18*(14), 1455–1458. doi:10.1097/WNR.0b013e3282ef6fc9

Piras, F., & Marangolo, P. (2010). When "Crack walnuts" lies in different brain regions: evidence from a voxel-based lesion-symptom mapping study. *Journal of the International Neuropsychological Society: JINS*, *16*(3), 433–442. doi:10.1017/S1355617710000068

Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition, 92*(1), 1-12. doi: 10.1016/j.cognition.2003.11.001

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *The Journal of Neuroscience, 32*(41), 14125–14131. doi: 10.1523/JNEUROSCI.3244- 12.2012.

Potts, M. B., Smith, J. S., Molinaro, A. M., & Berger, M. S. (2012). Natural history and surgical management of incidentally discovered low-grade gliomas. *Journal of Neurosurgery, 116*(2), 365-372. Retrieved from: http://thejns.org/doi/abs/10.3171/2011.9.JNS111068

Pradat-Diehl, P., Tessier, C., Vallat, C., Mailhan, L., Mazevet, D., Lauriot-Prevost, M. C., & Bergego, C. (2001). Conduction aphasia and phonemic disorder]. *Revue Neurologique*, *157*(10), 1245-1252. Retrieved from http://europepmc.org/abstract/med/11885517

Pringle, A. M., Taylor, R., & Whittle, J. R (1999). Anxiety and depression in patients with an intracranial neoplasm before and after tumor surgery. *British Journal of Neurosurgery, 13*(1), 46–51. doi. 10.1080/02688699944177

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–38). East Sussex, UK: Psychology Press Ltd.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological review*, *107*(3), 460. doi:10.1037/0894-4105.21.1.20

Raymer, A. M., Foundas, A. L., Maher, L. M., Greenwald, M. L., Morris, M., Rothi, L. J. G., & Heilman, K. M. (1997). Cognitive neuropsychological analysis and neuroanatomic correlates in a case of acute anomia. *Brain and language*, *58*(1), 137-156. doi:10.1006/brln.1997.1786

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 358–374. doi: 10.1016/S0022-5371(83)90236-0

Recht, L. D., Lew, R., & Smith, T. W. (1992). Suspected low-grade glioma: Is deferring treatment safe?. *Annals of Neurology, 31*(4), 431-436. doi: 10.1002/ana.410310413

Recht, L. D., McCarthy, K., O'Donnell, B. F., Cohen, R., & Drachman, D. A. (1989). Tumor-associated aphasia in left hemisphere primary brain tumors The importance of age and tumor grade. *Neurology, 39*(1), 48-48. doi: 10.1212/WNL.39.1.48

Reijneveld, J. C., Sitskoorn, M. M., Klein, M., Nuyen, J., & Taphoorn, M. J. B. (2001). Cognitive status and quality of life in patients with suspected versus proven low-grade gliomas. *Neurology, 56*(5), 618-623. doi: 10.1212/WNL.56.5.618

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Battery. Theory and clinical interpretation.* Neuropsychology Press, Tuscan, AZ.

Reverberi, C., Lavaroni, A., Gigli, G. L., Skrap, M., & Shallice, T. (2005). Specific impairments of rule induction in different frontal lobe subgroups. *Neuropsychologia*, *43*(3), 460-472. doi: 10.1016/j.neuropsychologia.2004.06.008

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443-447. doi: 10.1126/science.1100301

Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., & Carter, C. S. (2004b). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition, 56*(2), 129-140. doi: 10.1016/j.bandc.2004.09.016

Riva, D., & Devoti, M. (1996). Discontinuation of phenobarbital in children: effects on neurocognitive behavior. *Peadiatric Neurology, 14*(1), 36-40. doi: 10.1016/0887-8994(95)00224-3

Rizzo S., Venneri, A., & Papagno, C. (2002). A normative study of a famous face recognition and naming test. *Neurological Sciences, 23*(4), 153-159. doi: 10.1007/s100720200056

Robbins, T, W., James, M., Owen, A, M., Sahakian, B, J., Lawrence, A, D., McInnes, L., & Rabbit, P, M. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society*, *4*(5), 474-490. Retrieved from: https://www.cambridge.org/core/journals/journal-of-the-international-neuropsychological-society

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*(4), 374. doi: 10.1037/0096-3445.123.4.374

Robinson, G., Blair, J., & Cipolotti, L. (1998). Dynamic aphasia: an inability to select between competing verbal responses?. *Brain: a journal of neurology*, *121*(1), 77-89. doi: 10.1093/brain/121.1.77

Robinson, G., Shallice, T., Bozzali, M., & Cipolotti, L. (2010). Conceptual proposition selection and the LIFG: Neuropsychological evidence from a focal frontal group. *Neuropsychologia*, *48*(6), 1652-1663. doi:10.1016/j.neuropsychologia.2010.02.010

Robinson, G., Shallice, T., Bozzali, M., & Cipolotti, L. (2012). The differing roles of the frontal cortex in fluency tests. *Brain*, *135*(7), 2202-2214. doi: 10.1093/brain/aws142

Robinson, G., Shallice, T., Cipolotti, L. (2005). A failure of high level verbal response selection in progressive dynamic aphasia. *Cognitive Neuropsychology, 22*(6), 661-694. doi: 10.1080/02643290442000239

Robson, H., Sage, K., & Lambon Ralph, M. A. (2012). Wernicke's aphasia reflects a combination of acoustic-phonological and semantic control deficits: a case-series comparison of Wernicke's aphasia, semantic dementia and semantic aphasia. *Neuropsychologia*, *50*(2), 266-275. doi: 10.1016/j.neuropsychologia.2011.11.021

Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language, 72*, 193-218. doi: 10.1006/brln.1999.2285

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1-3), 107–142. doi:10.1016/0010-0277(92)90041-f

Roelofs, A. (2004). Comprehension-Based Versus Production-Internal Feedback in Planning Spoken Words: A Rejoinder to Rapp and Goldrick (2004). *Psychological Review*, *111*(2), 579–580. doi:10.1037/0033-295X.111.2.579

Rofes, A., & Miceli, G. (2014). Language mapping with verbs and sentences in awake surgery: a review. *Neuropsychology Review, 24*(2), 185-199. doi: 10.1007/s11065-014-9258-5

Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology*, *22*(7), 817–850. doi:10.1080/02643290442000365

Romani, C., Olson, A., Semenza, C., & Granà, A. (2002). Patterns of phonological errors as a function of a phonological versus an articulatory locus of impairment. *Cortex*, *38*(4), 541-567. doi:10.1016/s0010-9452(08)70022-4

Romero Lauro, J. R., Reis, J., Cohen, L. G., Cecchetto, C., & Papagno, C. (2010). A case for the involvement of phonological loop in sentence comprehension. Neuropsychologia, 48(14), 4003-4011. doi: 10.1016/j.neuropsychologia.2010.10.019

Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., & Karnath, H. O. (2012). Age-specific CT and MRI templates for spatial normalization. *Neuroimage*, *61*(4), 957-965. doi:10.1016/j.neuroimage.2012.03.020

Rorden, C., Karnath, H. O., & Bonilha, L. (2007). Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience*, *19*(7), 1081-1088. doi: 10.1162/jocn.2007.19.7.1081

Rueckert, L., & Grafman, J. (1996). Sustained attention deficits in patients with right frontal lesions. *Neuropsychologia, 34*(10), 953-963. doi: 10.1016/0028-3932(96)00016-4

Ruff, R. M., Allen, C. C., Farrow, C. E., Niemann, H., & Wylie, T. (1994). Figural fluency: differential impairment in patients with left versus right frontal lobe lesions. *Archives of Clinical Neuropsychology*, *9*(1), 41-55. doi: 10.1093/arclin/9.1.41

Saffran, E. M. (1990). Short-term memory impairment and language processing. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics: Advances in models of cognitive function and impairment* (pp. 137-168). Hillsdale, NJ: Lawrence Erlbaum Associates.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language, 37*(3), 440-479. doi: 10.1016/0093-934X(89)90030-8

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience*, *6*(1), 75-81. doi: 10.1038/nn987

Sanai, N., Mirzadeh, Z., & Berger, M. S. (2008). Functional outcome after language mapping for glioma resection. *The New England Journal of Medicine*, *358*(1), 18–27. doi:10.1056/NEJMoa067819

Santini, B., Talacchi, A., Squintani, G., Casagrande, F., Capasso, R., & Miceli, G. (2012). Cognitive outcome after awake surgery for tumors in language areas. *Journal of Neuro-Oncology*, *108*(2), 319-326. doi: 10.1007/s11060-012-0817-4

Santo Pietro, M. J., & Rigrodsky, S. (1986). Patterns of oral-verbal perseveration in adult aphasics. *Brain and Language*, *29*(1), 1-17. doi: 10.1016/0093-934X(86)90030-1

Santo-Pietro, M, J. & Rigrodsky, S. (1982). The effects of temporal and semantic conditions on the occurrence of the error response of perseveration in adult aphasics. *Journal of Speech and Hearing* Research, 25, 184–192. doi: 10.1044/jshr.2502.184

Sarter, M., Givens, B., & Bruno, J, P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain Research Reviews, 35*(2), 146-160. doi: 10.1016/S0165-0173(01)00044-3

Satoer, D., Kloet, A., Vincent, A., Clemens, D., & Visch-Brink, E. (2014). Dynamic aphasia following low-grade glioma surgery near the supplementary motor area: A selective spontaneous speech deficit. Neurocase: *The Neural Basis of Cognition, 20*(6), 704-716. doi: 10.1080/13554794.2013.841954

Satoer, D., Vincent, A., Smits, M., Dirven, C., & Visch-Brink, E. (2013). Spontaneous speech of patients with gliomas in eloquent areas before and early after surgery. *Acta neurochirurgica, 155*(4), 685-692. doi: 10.1007/s00701-013-1638-8

Schachar, R. J., Tannock, R., & Logan, G. (1993). Inhibitory control, impulsiveness, and attention deficit hyperactivity disorder. *Clinical Psychology Review*, *13*(8), 721-739. doi 10.1016/S0272-7358(05)80003-0

Schatz, J., Kramer, J. H., Ablin, A., & Matthay, K. K. (2000). Processing speed, working memory, and IQ: A developmental model of cognitive deficits following cranial radiation therapy. *Neuropsychology, 14*(2), 189-200. doi: 10.1037/0894-4105.14.2.189

Scheibel, R. S., Meyers, C. A., & Levin, V. A. (1996). Cognitive dysfunction following surgery for intracerebral glioma: influence of histopathology, lesion location, and treatment. *Journal of Neuro-Oncology, 30*(1), 61-67. doi: 10.1007/BF00177444

Schnur, T. T., Schwartz, M. F., Brecher, A. R., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227. doi:10.1016/j.jml.2005.10.002

Schroder, H. S., & Moser, J. S. (2014). Improving the study of error monitoring with consideration of behavioral performance measures. *Frontiers in Human Neuroscience*, *8*. doi 10.3389/fnhum.2014.00178

Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, *135*(12), 3799–3814. doi:10.1093/brain/aws300

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: voxel-based lesion-symptom mapping evidence from aphasia. *Brain*, *132*(12), 3411–3427. doi:10.1093/brain/awp284

Scott, R. M., & Wilshire, C. E. (2010). Lexical competition for production in a case of nonfluent aphasia: Converging evidence from four different tasks. *Cognitive Neuropsychology*, *27*(6), 505-538. doi:10.1080/02643294.2011.598853

Shafi, N., & Carozza, L. (2012). Treating cancer-related aphasia. *The ASHA Leader*. Retrieved from: https://www.researchgate.net/profile/Noel_Shafi/publication

Shalev, L., Ben-Simon, A., Mevorach, C., Cohen, Y., & Tsal, Y. (2011). Conjunctive Continuous Performance Task (CCPT) – A pure measure of sustained Attention. *Neuropsychologia, 49*(9), 2584-2591. doi: 10.1016/j.neuropsychologia.2011.05.006

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *298*(1089), 199-209. doi: 10.1098/rstb.1982.0082

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.

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

Shallice, T., & Skrap, M. (2011). Localisation through operation for brain tumour: a reply to Karnath and Steinbach. Cortex, 47(8), 1007-1009. doi: 10.1016/j.cortex.2010.12.006

Shallice, T., Stuss, D. T., Alexander, M. P., Picton, T. W., & Derkzen, D. (2008). The multiple dimensions of sustained attention. *Cortex, 44*(7), 794-805. doi: 10.1016/j.cortex.2007.04.002

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in psychology*, *5*(772), doi:10.3389/fpsyg.2014.00772

Shapiro, L. P., & Levine, B. A. (1990). Verb processing during sentence comprehension in aphasia. *Brain and Language, 38*(1), 21-47. doi: 10.1016/0093-934X(90)90100-U

Sheline, G. E., Wara, W. M., & Smith, V. (1980). Therapeutic irradiation and brain injury. *International Journal of Radiation, Oncology, Biology, Physics, 6*(9), 1215–28. doi: 10.1016/0360-3016(80)90175-3

Sloan Berndt, R., Haendiges, A. N., Mitchum, C. C., & Sandson, J. (1997b). Verb retrieval in aphasia. 2. Relationship to sentence processing. *Brain and Language*, *56*(1), 107-137. doi:10.1006/brln.1997.1728

Sloan Berndt, R., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997). Verb retrieval in aphasia. 1. Characterizing single word impairments. *Brain and Language*, *56*(1), 68- 106. doi:10.1006/brln.1997.1727

Smith, A. (1968). The symbol-digit modalities test: a neuropsychological test of learning and other cerebral disorders. *Learning disorders*, 83-91.

Smith, A. (1982) *Symbol Digit Modalities Test. Manual*. Western Psychological Services: Los Angeles

Smith, J. S., Chang, E. F., Lamborn, K. R., Chang, S. M., Prados, M. D., Cha, S., Tihan, T., VandenBerg, S., … & Berger, M. S. (2008). Role of extent of resection in the long-term outcome of low-grade hemispheric gliomas. Journal of Clinical Oncology, 26(8), 1338 – 1345. doi: 10.1200/JCO.2007.13.9337

Smith, D. B., Mattson, R. H., Cramer, J. A., Collins, J. F., Novelly, R. A., & Craft, B. (1987) Results of a nationwide Veterans Administration Cooperative Study comparing the efficacy and toxicity of carbamazepine, phenobarbital, phenytoin, and primidone. *Epilepsia, 28*(s3), S50–S58. doi: 10.1111/j.1528-1157.1987.tb05778.x

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 47*(2), 276. Doi: 10.1037/h0078826

Spreen, O. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press.

Stemberger, J. P. (1985). An interactive activation model of language production. *Progress in the psychology of language*, *1*, 143-186.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of Syntactic Comprehension by Positron Emission Tomography. *Brain and Language, 52*, 452–473. doi: 10.1006/brln.1996.0024

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of *Experimental Psychology, 18*(6), 643–662. doi:10.1037/h0054651

Stuss, D. T. (2011). Functions of the frontal lobes: relation to executive functions. *Journal of the international neuropsychological Society, 17*(5), 759-765. doi: 10.1016/j.neuropsychologia.2004.06.008

Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive syndrome?. Philosophical Transactions of the Royal Society of London: *Biological Sciences, 362*(1481), 901-915. doi: 10.1098/rstb.2007.2096

Stuss, D. T., Alexander, M. P., Hamer, L., Palumbo, C., Dempster, R., Binns, M., ... & Izukawa, D. (1998). The effects of focal anterior and posterior brain lesions on verbal fluency. *Journal of the International Neuropsychological Society*, *4*(03), 265-278. doi:10.1016/s0028-3932(97)00152-8

Stuss, D. T., Alexander, M. P., Palumbo, C. L., Buckle, L., Sayer, L., & Pogue, J. (1994). Organizational strategies with unilateral or bilateral frontal lobe injury in word learning tasks. *Neuropsychology*, *8*(3), 355. doi: 10.1037/0894-4105.8.3.355

Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., ... & Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia, 43*(3), 396-417. doi: 10.1016/j.neuropsychologia.2004.06.010

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*(8), 771-786. doi: 10.1016/S0028-3932(01)00013-6

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. A*nnals of the New York Academy of Sciences, 769*(1), 191-212. doi: 0.1111/j.1749-6632.1995.tb38140.x

Suchy, Y., Kraybill, M. L., & Larson, J. C. G. (2010). Understanding design fluency: Motor and executive contributions. *Journal of the International Neuropsychological Society*, *16*(1), 26-37. Doi: 10.1017/S1355617709990804

Swick, D., Ashley, V., & Turken, U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC neuroscience*, *9*(1), 102. doi: 10.1186/1471-2202-9-102

Talacchi, A., Santini, B., Savazzi, S., & Gerosa, M. (2011). Cognitive effects of tumour and surgical treatment in glioma patients. *Journal of Neuro-Oncology*, *103*(3), 541– 549. doi:10.1007/s11060-010-0417-0

Tandon, P., & Mahapatra, A. K. (1993). Operations on gliomas involving speech centres. In Pasztor E., Vajda, J., Loew, F (eds) *Language and Speech* (pp. 67-71), vol 56. Springer: Vienna.

Taphoorn, M. J. (2003,). Neurocognitive sequelae in the treatment of low-grade gliomas. In *Seminars in Oncology*, *30*, 45-48. doi: 10.1053/j.seminoncol.2003.11.023

Taphoorn, M. J. B., & Klein, M. (2004). Cognitive deficits in adult patients with brain tumours. *The Lancet*, *3*(3), 159–168. doi:10.1016/S1474-4422(04)00680-5

Team, A. D. (2008). Audacity (Version 1.2. 6)[Computer software]. Available: audacity. sourceforge. net/download.

Teixidor, P., Gatignol, P., Leroy, M., Masuet-Aumatell, C., Capelle, L., & Duffau, H. (2007). Assessment of verbal working memory before and after surgery for low-grade glioma. *Journal of Neuro-Oncology, 81*(3), 305–13. doi: 10.1007/s11060-006-9233-y

Thomas, R., O'Connor, A. M., & Ashley, S. (1995). Speech and language disorders in patients with high grade glioma and its influence on prognosis. *Journal of Neuro-Oncology*, *23*(3), 265–270. doi:10.1007/bf01059960

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, *23*(3), 513–522. doi:10.1016/s0896-6273(00)80804-1

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A re-evaluation.

Proceedings of the National Academy of Sciences of the United States of America, *94*(26), 14792–14797. doi:10.1073/pnas.94.26.14792

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences, 95*(26), 15855-15860. doi: 10.1073/pnas.95.26.15855

Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*(4-5), 473-490. doi: 10.1080/02687039708248485

Thompson, S. A., Patterson, K., & Hodges, J. R. (2003). Left/right asymmetry of atrophy in semantic dementia Behavioral–cognitive implications. *Neurology*, *61*(9), 1196-1203. doi:10.1212/01.wnl.0000091868.28557.b8

Tremblay, P., & Gracco, V. L. (2006). Contribution of the frontal lobe to externally and internally specified verbal responses: fMRI evidence. *Neuroimage*, *33*(3), 947-957. doi: 10.1016/j.neuroimage.2006.07.041

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language, 39*(1), 102-123. doi: 10.1006/jmla.1998.2565

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory & Language, 33*(3), 285–318. doi: 10.1006/jmla.1994.1014

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*(7), 1790–1800. doi: 10.1016/j.cortex.2012.10.014

Tucha, O., Smely, C., Preier, M., & Lange, K. W. (2000). Cognitive deficits before treatment among patients with brain tumors. *Neurosurgery*, *47*(2), 324-334.doi: 10.1097/00006123-200008000-00011

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 151-161. doi:10.1016/s0022-5371(84)90104

Vallarb, G., Di Bettac, A. M., & Silveri, M. C. (1997). The phonological short-term store-rehearsal system: patterns of impairment and neural correlates. *Neuropsychologia, 35*(6), 795-812. doi: 10.1016/S0028-3932(96)00127-3

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*(3), 443 – 466. doi: 10.1037/0278-7393.31.3.443

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., ... & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage*, *30*(4), 1414-1432. doi: 10.1016/j.neuroimage.2005.11.002

Wacker, A., Holder, M., Will, B. E., Winkler, P. A., & Ilmberger, J. (2002). Vergleich von Aachener Aphasie-Test, klinischer Untersuchung und Aachener Aphasie-Bedside Test bei Hirntumorpatienten. *Der Nervenarzt, 73*(8), 765–769. doi: 10.1007/s00115- 002-1358-4

Wagner, S., Sebastian, A., Lieb, K., Tüscher, O., & Tadić, A. (2014). A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects. *BMC Neuroscience*, *15*(19). doi:10.1186/1471-2202-15-19

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433-441. doi: 10.1518/001872008X312152

Wechsler D: Wechsler adult intelligence scale-revised. San Antonio: Psychological Corporation, 1981

Wechsler D. *Wechsler Adult Intelligence Scale, III*. San Antonio, TX: The Psychological Corporation, 1997.

Welzel, G., Steinvorth, S., & Wenz, F. (2005). Cognitive effects of chemotherapy and/or cranial irradiation in adults. *Strahlentherapie und Onkologie*, *181*(3), 141-156. doi: 10.1007/s00066-005-1351-5

Wernicke, C. (1874). *Der aphasische Symptomencomplex*. Breslau: Max. Cohn & Weigert.

Whittle, I. R., Pringle, A. M., & Taylor, R. (1998). Effects of resective surgery for left-sided intracranial tumours on language function: A prospective study. *Lancet, 351*(9108), 1014–1018. doi:10.1016/s0140-6736(97)08295-0

Whittle, I. R., Smith, C., Navoo, P., & Collie, D. (2004). Meningiomas. *The Lancet*, *363*(9420), 1535-1543. doi: 10.1016/S0140-6736(04)16153-9

Wilkins, A. J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia*, *25*(2), 359-365. doi: 10.1016/0028-3932(87)90024-8

Wilshire, C. E. (2002). Where do aphasic phonological errors come from? Evidence from phoneme movement errors in picture naming. *Aphasiology*, *16*(1-2), 169-197. doi:10.1080/02687040143000528

Wilshire, C. E. (2014) *Cognitive Neuropsychology: Exploring the Mind Through Brain Dysfunction* (in prep).

Wilshire, C. E., & McCarthy, R. A. (1996). Experimental investigations of an impairment in phonological encoding. *Cognitive Neuropsychology*, *13*(7), 1059-1098. doi:10.1080/026432996381782

Wilshire, C. E., & McCarthy, R. A. (2002). Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology, 19*(2), 165-186. doi: 10.1080/02643290143000169

Wilshire, C. E., Keall, L. M., Stuart, E. J., & O'Donnell, D. J. (2007). Exploring the dynamics of aphasic word production using the picture–word interference task: A case study. *Neuropsychologia*, *45*(5), 939-953. doi:10.1016/j.bandl.2005.07.057

Wilshire, C. E., Lukkien, C. C., & Burmester, B. R. (2014). The sentence production test for aphasia. *Aphasiology, 28*(6), 658-691. doi: 10.1080/02687038.2014.893555

Wu, A. S., Witgert, M. E., Lang, F. F., Xiao, L., Bekele, B. N., Meyers, C. A., ... & Wefel, J. S. (2011). Neurocognitive function before and after surgery for insular gliomas: Clinical article. *Journal of Neurosurgery*, *115*(6), 1115-1125. doi: 10.3171/2011.8.JNS11488.

Yamadori, A. (1981). Verbal perseveration in aphasia. *Neuropsychologia*, *19*(4), 591-594. doi: 10.1016/0028-3932(81)90026-9

Ziegler, W., Kilian, B., & Deger, K. (1997). The role of the left mesial frontal cortex in fluent speech: Evidence from a case of left supplementary motor area haemorrhage. *Neuropsychologia, 35*(9), 1197-1208. doi: 10.1016/S0028-3932(97)00040-7

Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language, 39*(1), 14-32. doi: 10.1016/0093-934X(90)90002-X