Development of Systematic Behavioural Observation to Quantify Ongoing Cognitive Activity Limitations after Brain Injury

Mark W Lewis

2010

Abstract

One of the goals of cognitive rehabilitation following traumatic brain injury is to help people perform everyday tasks. However, options for the rigorous assessment of everyday cognitive effectiveness after rehabilitation are limited. Performance on neuropsychological tests is only moderately correlated with everyday functioning, while previous measures of everyday functioning include only fairly general estimates of overall cognitive functioning. The aim of the current study was to develop an ecologically valid measure that captured a number of subdomains of executive functioning, using systematic behavioural observation of an everyday task. The initial phase of the research involved identifying an everyday task that was sufficiently complex to ensure that executive functioning was utilised in the completion of the task. Participants with traumatic brain injury were then asked to prepare chocolate brownies, using a recipe provided, and a hot drink. Participants were allowed to use any compensatory strategy to help complete the task. Participant performance was directly observed by an examiner and videotaped for subsequent inter-rater reliability. Two independent raters assessed nine components of executive functioning. During this phase, the examiner manuals were modified improving inter-rater reliability. The final version of the measure was then trialled with participants with and without traumatic brain injury. Final inter-rater reliability indicated the approach had merit. Significant and moderate correlations were found between traditional measures of executive functioning and the everyday task. This study employed systematic behavioural observation to obtain fine-grained information regarding a person's cognitive functioning. With further development, this approach may prove useful for targeting and monitoring specific functional difficulties during cognitive rehabilitation.

Preface

I would like to acknowledge Dr Duncan Babbage, whose continual input, guidance, and assistance was critical in the completion of this study. I also acknowledge Professor Janet Leathem, whose insightful comments and final review were much appreciated. To those who participated in the study and to the organisations who kindly advertised the study and directed clients to the study (i.e., the Brain Injury Association of New Zealand and CAVIT ABI Rehabilitation), I thank you. Finally, I would like to acknowledge Celia Falchi for assisting in rating participant performance and Belinda Snow for performing the early renditions of the functional task.

Table of Contents

Abstract ii

Preface iii

Table of Contents iv

List of Tables v

List of Figures vii

List of Abbreviations viii

Chapter One: Introduction 1

Chapter Two: Pre-Study Survey 33

Chapter Three: Developmental Trial 50

Chapter Four: Main Assessment 72

References 118

Appendices 140

List of Tables

Table 1. Aspects of Commonality Between Five Models of EF 7

Table 2. Demographic Characteristics of Participants in the Pre-Study Survey 35

Table 3. Comparisons Between Participants and Family Members on the EBIQ Subscales 42

Table 4. Comparison Between Ability Scores on the Activities of Daily Life Questionnaire as Reported by Participants and Family Members 43

Table 5. Comparison Between Occurrence of Activities of Daily Life Questionnaire Items as Reported by Participants and Family Members 44

Table 6. Number of Participants Who Use Compensatory Strategies for Activities of Daily Life Questionnaire Items (n = 13) 45

Table 7. Demographic Characteristics of Participants in the Developmental Trail 52

Table 8. Demographic Characteristics of Participants in the Main Assessment 74

Table 9. Pearson's Correlations Between the Functional Task and the Sorting and Tower Tests for Participants with TBI (n = 27) 89

Table 10. Correlations Between the Functional Task and Self-Reported EBIQ Scores for Participants with TBI (n = 26) 90

Table 11. Correlations Between the Functional Task and Family Member Reported EBIQ Scores for Participants with TBI (n = 15) 91

Table 12. Comparisons Between Self-Rating and Family Member-Rating for Participants with TBI on the EBIQ Subscales 92

Table 13. Correlation Matrix Between the Functional Task and the Sorting and Tower Tests for All Participants (n = 43) 93

Table 14. Correlations Between the Functional Task and Self-Reported EBIQ Scores for All Participants (n = 41) 94

Table 15. Comparison Between the Functional Ratings for the TBI and Normative Groups 95

Table 16. Comparison Between EBIQ Scores for the TBI and Normative Groups 96

Table 17. Multiple Regression Analysis of Variables in Relation to the Functional Task (n = 43) 97

List of Figures

Figure 1. Interactions between components of the ICF 16

Figure 2. Relationship between the different phases of the study. 32

Figure 3. Decision tree for awarding functional ratings for planning and organisation, executive memory, initiation, cognitive shifting, impulsivity, directed and sustained attention, and time management 60

Figure 4. Decision tree for awarding functional ratings for error detection and error correction 62

List of Abbreviations

TBI — Traumatic brain injury

LOC — Loss of consciousness

PTA — Post-traumatic amnesia

DAI — Diffuse axonal injury

EF — Executive functioning

WCST — Wisconsin Card Sorting Test

ACC — Anterior cingulate cortex

PFC — Prefrontal cortex

ADL — Activities of daily life

ICF — International Classification of Functioning, Disability, and Health

BADS — Behavioural Assessment of Dysexecutive Syndrome

FIM — Functional Independence Measure

FAM — Functional Assessment Measure

MET-HV — Multiple Errands Test, Hospital Version

TOFEA — Test of Functional Executive Abilities

AMPS — Assessment of Motor and Process Skills

EBIQ — European Brain Injury Questionnaire

D-KEFS — Delis-Kaplin Executive Function System

ICC — Intraclass correlation coefficient

Chapter One

Introduction

A traumatic brain injury (TBI) refers to brain injuries that occur when external forces act on the head. As a result of these injuries, confusion, unconsciousness, amnesia, or other neurological conditions may occur (Lucas, 1998). After a TBI, people will likely experience a range of cognitive difficulties. Particularly vulnerable to TBI are the cognitive processes known as executive functions, which refer to the higher-level cognitive processes typically associated with the frontal lobes. These processes, which include abstraction, organisation, planning, and insight (World Health Organization, 2001), preside over other cognitive processes, such as memory and attention, in solving everyday problems (McCullagh & Feinstein, 2005).

Formal neuropsychological assessment is a useful tool for quantifying cognitive functioning after someone has had a brain injury as it determines strengths as well as deficits. Once this information in known, strategies can be taught that will help compensate for, or treat, any difficulties (Wilson, 1996). Although it has been generally assumed that formal neuropsychological tests are ecologically valid, research shows only modest correlations; Pearson's correlations range from .24 to .50, between neuropsychological tests and everyday behaviours (Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Chaytor, Schmitter-Edgecombe, & Burr, 2006; Manchester, Priestley, & Jackson, 2004). Thus, treatment strategies that are based on information gathered from traditional neuropsychological tests might not be useful in dealing with everyday problems. Furthermore, formal neuropsychological assessment highlights the cognitive *in*abilities people display, rather than highlighting their cognitive abilities (Wilson, 1996).

To date, there are few neuropsychological measures that focus on everyday cognitive abilities and fewer still that utilise behavioural observation. As Wilson (1996) explains, neuropsychological assessments should incorporate real-world tasks in order for the assessment to more fully inform the rehabilitation process. The current study is an attempt to develop and preliminarily trial a systematic behavioural observation approach in assessing the ability to perform a real-world task that taps executive functions.

In preparing for the study, several issues were initially considered. First, it was necessary to understand the nature of traumatic brain injury and the organic changes that occur as a result of insult to the brain. Second, clarity regarding what executive functioning is and why these cognitive processes should be studied was undertaken. Third, a comparison between how cognitive processes are assessed and how they are treated, with specific reference to the International Classification of Functioning, Disability, and Health (World Health Organization, 2001), was made. Finally, a review of current assessment measures that utilise real-world tasks that tap executive functioning was conducted.

Neuropathology of Traumatic Brain Injury

A TBI can be classified into a number of different categories based on the severity, timing, type, and location of the injury. The severity of the TBI a person receives is usually assessed by the length of loss of consciousness (LOC), post-traumatic amnesia (PTA), and scores on the Glasgow Coma Scale. A TBI is classified as mild if the LOC is less than 30 minutes, whereas a TBI is classified as moderate to severe if LOC is greater than 30 minutes. Generally speaking, the longer the LOC, the greater the severity. A TBI is similarly classified as mild if the PTA a person experiences is less than 24 hours, as moderate if the duration of PTA is between one to six days, and as severe if the duration of PTA is seven or more days (Lucas, 1998). Scores on the Glasgow Coma Scale (Teasdale & Jennett, 1974) also indicate severity of TBI, where a score of 13-15 is classified as mild, 9-12 as moderate, and 3-8 as severe (Teasdale, Murray, Parker, & Jennett, 1979).

The time at which TBI occurs can be divided into two categories. Primary damage occurs at the point that external forces act on the head and can lead to contusions, lacerations, or diffuse axonal injury (DAI). Secondary damage occurs as a result of processes initiated during the primary damage (e.g., further DAI, swelling, or infection; Graham, 1995). For example, primary damage occurs when, during a TBI, axons become elongated and ripple, causing the axonal cytoskeleton to become damaged (Smith, Meaney, & Shull, 2003). As a result of the damaged cytoskeleton, secondary damage occurs through a chain reaction. "The damaged axolemma allows the influx of normally excluded extracellular ions, such as calcium, that in turn activate the cysteine proteases, calpain and caspase, to disrupt the cytoskeleton" (Povlishock & Katz, 2005, p. 406). This results in swollen, damaged axons that impede the neuronal impulse reaching the presynaptic terminal; DAI (Smith et al.).

A TBI can be further classified based on how the brain is assaulted. In a closed TBI, the brain contacts the skull causing damage to the peaks of the gyri (Graham, 1995; Halliday, 1999). The site of the damage depends on the movement of the skull and the movement of the object that contacts the skull. A coup injury, or damage to the area at the site of the impact, results when a stationary skull is struck by a moving object. A contre-coup injury, or damage to the area that is opposite the site of the impact, results when a moving skull strikes a stationary object (Shaw, 2002; Whiting & Zernicke, 1998). Coup and contre-coup injuries can occur independently or concurrently. In an open TBI, the skull and meninges are penetrated and the brain is directly assaulted. The frontal lobes, and to a lesser extent, the temporal lobes, are sensitive to coup and contre-coup injury as there are numerous bony protrusions on the inside of the skull and are sensitive to open TBI because of their relative size compared to the rest of the brain (McDonald, Flashman, & Saykin, 2002).

The location of structural damage can also be used to classify TBI. Focal injury refers to damage to a specific area and can present as contusions, lacerations, intracranial haematoma, or increased intracranial pressure. Diffuse injury refers to damage that is extensive, although the term also refers to damage that is multifocal. DAI, hypoxia, ischemia, and infection are all examples of diffuse brain injury (Gennarelli & Graham, 2005).

Following the initial haemorrhaging that may occur during a TBI, the brain heals, leaving depressed scars (Graham, 1995). The plasticity of the central nervous system allows some recovery to the damaged areas through neural regeneration, sprouting, and synaptogenesis. However, scars impede the regeneration of pre-injury neural connections (Kolb & Cioe, 2004) and the number of neurons that do regenerate is estimated to be less than 10% (Dusart et al., 2005). Axonal sprouting, despite being a well established phenomenon (Hagg, 2006), may not regenerate all of the pre-injury connections. Instead, the new connections may become maladaptive (Constantinidou, Thomas, & Best, 2004). Finally, the synaptogenesis that occurs following TBI does not necessarily result in a return to pre-injury status (Scheff et al., 2005). The ability to achieve functional recovery following TBI was summed up by Kolb (2004) who stated that although not impossible, a return to pre-injury ability is most unlikely.

The incidence rates of TBI in New Zealand are difficult to determine as many people who receive a mild TBI do not report to hospital or to their general practitioner

(New Zealand Guidelines Group, 2006). However, estimates of the incidence of TBI in New Zealand can be made. For every 100 people who report to hospital with mild TBI, it is estimated that an additional 60 people are treated independently by their general practitioner. This suggests that the incidence of people over the age of 15 with TBI, who come to the attention of medical professionals, is 700 per 100,000 (New Zealand Guidelines Group, 2006). Wrightson and Gronwall (1998) had elsewhere reported a hospital attendance rate for mild TBI of 437 per 100,000 for people over the age of 15, and a rate for mild TBI of 252 per 100,000 for people under the age of 15. These rates, which are similar to international rates of TBI (Cassidy et al., 2004), suggest that approximately 20,000 to 30,000 people sustain a TBI each year in New Zealand (New Zealand Guidelines Group).

Defining Executive Functioning

The nature and extent of cognitive deficits experienced following TBI depend on a number of different factors, including the extent of DAI, the location and size of the lesions (Katz & Alexander, 1994), age at injury, and pre-existing or concurrent conditions (McCullagh & Feinstein, 2005). Despite these mitigating factors, memory, learning, attention, and executive function deficits are common following TBI, and are universal following severe TBI (Brooks, 1990; McCullagh & Feinstein). Of these cognitive deficits, there is a particular need to understand executive functioning (EF) as it influences a host of cognitive processes and people who attend neuropsychological services are more likely to display executive dysfunction than any other cognitive deficit (Stuss & Levine, 2002).

EF have been defined as the ability to generate and execute a choice (Malloy, Cohen, & Jenkins, 1998), to problem solve and complete goal-directed actions (Morse

& Montgomery, 1992), and to react in an appropriate manner to unfamiliar stimuli (Lezak, Howieson, Loring, Hannay, & Fischer, 2004). One reason for the different definitions of EF espoused by different researchers is that the definition of EF is a theoretical one, rather than an operational one (Burgess, 1997). People with executive dysfunction may display intact EF in some situations and impaired EF in other situations or may display some aspects of EF that are intact and other aspects that are impaired. This is unlike a person with amnesia, who will display dense memory impairments in every situation, thus leading easily to an operational definition of amnesia (Burgess). An operational definition of EF is further hampered by the fact that they do not refer to a single executive function (Stuss & Alexander, 2000), but to a group of different cognitive processes that work together to perform a particular objective (Burgess & Simons, 2005). Such is the nature of EF that in an attempt to define EF, some researchers appear to essentially list various cognitive processes (see Sohlberg & Mateer, 2001).

Models of EF can be developed through different methods. In this section, two models that were based on theories of the frontal lobe will be discussed, followed by three models that were based on factor analysis (see Table 1). Each of these models stress different aspects of EF, thus highlighting the difficulty researchers have in studying EF.

Table 1

Aspects of Commonality Between Five Models of EF

Aspects	Lezak et al. (2004)	Sohlberg and Mateer (2001)	Boone et al. (1998)	Burgess et al. (1998)	Busch et al. (2005)
1	Volition	Initiation and drive			PF/CF
2	Planning	Organisation		Intentionality, Executive memory	PF/CF
3	Purposive action			Intentionality	
4	Effective performance	Awareness	Cognitive flexibility		
5		Response inhibition		Inhibition	
6		Task persistence	Basic/ divided attention	Intentionality	PF/CF, Mental control
7		Generative thinking		Inhibition	PF/CF
8			Speeded processing		
9				Executive memory	Memory errors
10				Positive affect	
11				Negative affect	

Note. PF/CF = productive fluency/cognitive flexibility

According to Lezak et al. (2004), EF is comprised of four cognitive processes: volition, planning, purposive action, and effective performance. Each of these processes, although distinct, work together to produce goal-directed behaviour. Volition, or intentionality, is the process of forming intentions so that one's wants and needs can be accomplished. In order for this to occur, people must be able to conceptualise their current situation and compare that conceptualisation to a hypothetical future where their wants and needs are met. A critical component to volition is having the motivation to initiate and maintain certain behaviours that will lead to the accomplishment of one's goals. Poor insight, motivation, or initiation can impair a person's ability to utilise volition (McDonald et al., 2002). If volition is impaired, goal-directed behaviour is impaired because people simply do nothing (Lezak et al.).

Planning refers to the process involved in identifying and arranging the steps necessary to perform goal-directed behaviour. Associated with planning is the ability to anticipate the future, with respect to one's own behaviour and environmental changes, in order to increase the likelihood that the goal will be accomplished. Furthermore, planning requires that people have intact memory processes, impulse control, the ability to sustain attention, and be motivated to accomplish a goal. If one or more of these aspects of planning are impaired, the plan a person creates will likely fail. A plan will similarly fail if a person's goals are unrealistic (Lezak et al., 2004).

After developing a plan to accomplish a goal, the plan needs to be implemented. The initiation and monitoring of actual behaviour, in relation to planned behaviour, is referred to as purposive action. A person with impaired purposive action may have the intention, time, and resources to perform a particular behaviour, but never actually perform the behaviour (Lezak et al., 2004). Furthermore, goal-directed behaviours may fail at the purposive action stage due to impaired social awareness or a lack of insight into one's own behaviour (McDonald et al., 2002).

Finally, effective performance refers to the qualitative aspects of the behaviour. Components of effective performance include monitoring and self-correction (Lezak et al., 2004). A goal-directed behaviour can fail at this stage if people have poor insight into their performance, do not correct mistakes because they have difficulty with cognitive shifting, or do not correct mistakes because they are unmotivated to self-correct (McDonald et al., 2002).

Sohlberg and Mateer (2001) proposed a different model of EF that incorporates six components: initiation and drive, response inhibition, task persistence, organisation, generative thinking, and awareness. Initiation and drive, which is similar to Lezak et al.'s (2004) volition, refer to the ability to start a goal-directed behaviour. If this aspect of EF is impaired, people will be apathetic and unable to initiate intended actions. Response inhibition refers to the ability to stop a goal-directed behaviour and to inhibit impulsive behaviours. Failure in this aspect of EF results in impulsive, perseverative responding. Task persistence refers to the ability to sustain attention until the goal is complete. A critical component to task persistence is working memory. If a person displays task persistence impairments, they will likely discontinue their goal-directed behaviour before the goal has been accomplished (Sohlberg & Mateer).

Organisation, which is similar to Lezak et al.'s (2004) planning, refers to the ability to identify a particular goal and organise the steps necessary to accomplish that goal. People with impaired organisation will likely fail to complete their goals because the identified steps were not sequenced properly. Generative thinking denotes the ability to organise information in an effective manner and is related to how effectively people can find solutions to their problems. People who are creative and are cognitively flexible are more likely to achieve their goals than are people who are cognitively rigid. The final component of EF as described by Sohlberg and Mateer (2001) is awareness and refers to the ability to monitor and adjust behaviour to match one's planned strategy and to match personal and environmental changes. Insight is critical for successful goal-directed actions in this component of EF. This component is similar to Lezak et al.'s effective performance.

The models proposed by Lezak et al. (2004) and Sohlberg and Mateer (2001) are similar to models of EF that were derived using factor analysis. For example, Boone, Pontón, Gorsuch, González, and Miller (1998) developed a three factor model of EF based on a factor analysis of formal neuropsychological tests. The first factor was based on scores on the Wisconsin Card Sorting Test (WCST; Heaton, 1981) and was termed cognitive flexibility. The cognitive aspects of this factor were the ability to adjust behaviour in response to environmental stimuli, problem-solve, and shift cognitively (Boone et al.). This factor was similar to Lezak et al.'s effective performance and Sohlberg and Mateer's awareness. Speeded processing was the second factor and included scores on the Stroop Color-Word Test (Comalli, Wapner, & Werner, 1962) and the FAS test (Miller, 1984). The cognitive aspects of the second factor included information processing speed, rather than response inhibition and cognitive productivity that these tests are typically associated with (Boone et al.). The final factor, which is similar to Sohlberg and Mateer's task persistence, was termed basic/divided attention and short-term memory and included scores on the Auditory Consonant Trigrams (Boone, Miller, Lesser, Hill, & D'Elia, 1990), Digit Span and Digit Symbol from a modified version of the Wechsler Adult Intelligence Scale - Revised (Adams, Smigielski, & Jenkins, 1984), and the Rey-Osterreith Complex Figure (Boone, Lesser, Hill-Gutierrez, Berman, & D'Elia, 1993). The cognitive aspects of this final factor included basic and divided attention and short-term memory (Boone et al.).

Burgess et al. (1998) developed a five factor model of EF based on a factor analysis of the Dysexecutive Questionnaire (Burgess, Alderman, Wilson, Evans, & Emslie, 1996). The five factors were then correlated with formal neuropsychological tests. Factor one, labelled inhibition, included questions about impulsivity, disinhibition, and abstract reasoning. This factor shares aspects with Sohlberg and Mateer's (2001) response inhibition and generative thinking. The formal tests of EF that correlated with factor one were Trail Making (Armitage, 1946), FAS, Animal Naming (McKenna, Mortimer, & Hodges, 1994), the Six Element Test (Burgess, Alderman, Evans et al., 1996), and the Cognitive Estimates Test (Shallice & Evans, 1978). Along with these tests, full-scale IQ as measured by the Wechsler Adult Intelligence Scale - Revised (WAIS-R; Wechsler, 1981) and the National Adult Reading Test (Nelson & Willison, 1991) correlated significantly with factor one. Interpreting factor one is difficult given that measures of general intelligence correlated with a factor supposedly capturing some aspect of EF. However, it might be that this factor represents a kind of general EF and impairments in inhibition lead to decreased intelligence scores (Burgess et al.).

Burgess et al.'s (1998) factor two, labelled intentionality, included questions relating to goal planning and maintenance, insight, and distractibility. This factor shares aspects with Lezak et al.'s (2004) planning and purposive action and with Sohlberg and Mateer's (2001) organisation and task persistence. The only executive test that correlated significantly with factor two was the Six Elements Test. This was expected given that the Six Elements Test assesses the ability to create, initiate, and maintain intentions (Burgess, 1997). Factor three, labelled executive memory, included questions relating to sequencing, confabulation, and perseveration (Burgess et al.), aspects of which can be found in Lezak et al.'s planning and Sohlberg and Mateer's organisation. Although the WCST has been used as a general measure of EF (Demakis, 2003), the Modified WCST (Nelson, 1976) only correlated with factor three. This suggests that the Modified WCST taps specific aspects of EF. Factor three also correlated with various memory tests, suggesting that this factor taps EF as they relate to memory. Factors four and five tapped positive and negative affect, respectively. No formal neuropsychological test correlated with factor four and only one correlated, moderately, with factor five (Burgess et al.).

One of the main criticisms of Boone et al.'s (1998) and Burgess et al.'s (1998) research is that participants in their studies displayed mixed aetiologies (Busch, McBride, Curtiss, & Vanderploeg, 2005). It might be that the factors found reflect differences in EF, however, the factors may also reflect the specific effect of the different aetiologies. In an attempt to avoid this criticism, Busch et al. recruited participants with a single aetiology, non-penetrating traumatic brain injury. They developed a model of EF based on three factors. There were two components in the first factor, labelled productive fluency/cognitive flexibility. Measures of initiation, sustained attention, and self-generated responding loaded onto the first component, whereas measures of cognitive flexibility, shifting, and sequencing loaded onto the second component. Factor one shares aspects with Lezak et al.'s (2004) volition and planning and with Sohlberg and Mateer's (2001) initiation and drive, organisation, task persistence, and generative thinking. Factor two, similar to Sohlberg and Mateer's task persistence, was labelled mental control and was associated with measures sensitive to distraction, working memory, and sustained attention. Variables from the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987) and the Visual Spatial Learning Test (Malec, Ivnik, & Hinkeldey, 1991) loaded onto the third factor and represented reporting erroneous memories (Busch et al.).

The differing definitions and models of EF should not imply that the construct of EF is not a unified one. When EF are compared to non-EF, a distinction arises (Alvarez & Emory, 2006). EF represent high-level cognitive processes, whereas non-EF represent lower-level, subordinate, cognitive processes (Bennett, Ong, & Ponsford, 2005;

McCullagh & Feinstein, 2005; Stuss & Levine, 2002). EF can also be conceptualised as cognitive processes involved in coordinating neural activity, whereas non-EF represent those processes involved in processing neural activity and sensory stimuli (Alvarez & Emory).

Neurocircuitry of Executive Functioning

The frontal lobes comprise roughly 30% of the cortex of the entire brain (Goldman-Rakic, 1984) and have long been recognised as critical for EF (McDonald et al., 2002). Such has been the link between EF and the frontal lobes that the term executive dysfunction and frontal lobe syndrome have been used interchangeably (Elliott, 2003). Many of the early studies into EF examined the behaviour of people with brain lesions. For instance, Milner (1963) found that people who had undergone a frontal lobectomy displayed more perseverations and attained fewer categories on the WSCT than did people who had undergone a posterior lobectomy. In their meta-analysis, Alvarez and Emory (2006) found an additional 10 studies indicating that people with frontal lesions performed worse on the WSCT than did people with posterior lesions. Despite these findings, Alvarez and Emory also found four studies in which people with frontal lesions performed similarly on the WSCT to people with posterior lesions. These results have been found by other research (see Demakis, 2003; van den Broek, Bradshaw, & Szabadi, 1993). Neuroimaging studies have supported lesions studies in showing that along with the frontal regions, posterior areas are also activated during tasks sensitive to EF (Alvarez & Emory).

The view that EF can be attributed to either frontal, posterior, or a combination of both locations of the brain, however, is rather simplistic and does not take into account various neural pathways that have been found to influence functioning (Heyder, Suchan,

& Daum, 2004). For example, the anterior cingulate cortex (ACC) has been found to be active, along with lateral areas of the prefrontal cortex (PFC), during tasks sensitive to response inhibition (Braver, Barch, Gray, Molfese, & Snyder, 2001) and self-monitoring (MacDonald, Cohen, Stenger, & Carter, 2000). This has led researchers to conclude that an ACC-PFC pathway is active during EF (Haxby, Petit, Ungerleider, & Courtney, 2000). In addition to the ACC-PFC pathway, multiple cortico-subcortical pathways are involved in EF (Heyder et al.).

Three of the five neural pathways that link the basal ganglia to the frontal cortex, the dorsolateral PFC, orbitofrontal PFC, and ACC, are believed to be involved in EF (Alexander, DeLong, & Strick, 1986). The orbitofrontal PFC-basal ganglia pathway has been found to be activated when, during the WCST, people change their response set due to comments made by the examiner (Monchi, Petrides, Petre, Worsley, & Dagher, 2001). Despite some imaging studies, most of the evidence for the role of these pathways in EF comes from studies into Parkinson’s and Huntington’s diseases (Heyder et al., 2004). People with these degenerative diseases typically show impairments in EF, which has been ascribed to damage to the basal ganglia-dorsolateral PFC pathway (Stocchi & Brusa, 2000).

A pathway linking the PFC to the cerebellum has also been found to be involved in EF (Heyder et al., 2004; Kim, Uğurbil, & Strick, 1994). There is mixed evidence regarding the performance of people with cerebellar dysfunction on the WCST. Most of the evidence suggests there is no difference between the performance of people with cerebellar dyfunction and controls (Daum, Snitz, & Ackermann, 2001). However, Riva and Giorgi (2000) and Bürk, Daum, and Rüb (2006) found that people with cerebellar dyfunction displayed more perseverative behaviour than did age-matched controls. The finding that EF are impaired following damage to the cerebellum has even led

Schmahmann and Sherman (1997) to create the term "cerebellar cognitive affective syndrome" (p. 438).

In summary, a variety of intact brain regions and pathways are necessary for effective EF. Impairments in EF are likely following coup and contra coup injuries to the cerebral cortex (McDonald et al., 2002). Similarly, people may display executive deficits following DAI in the basal ganglia and cerebellum (Heyder et al., 2004; Parizel et al., 1998). The diffuse nature of the neurocircuitry of EF may explain why executive dysfunction is a common complaint following TBI (Stuss & Levine, 2002).

International Classification of Functioning, Disability, and Health

Cognitive rehabilitation following TBI has traditionally been based on two assumptions: (a) that there is a linear relationship between body functions and structures and the ability to perform activities of daily life (ADL; New Zealand Guidelines Group, 2006); and (b) improvements made in rehabilitation settings will generalise to everyday settings. These assumptions characterise the medical model and its focus on the individual. However, this model does not explain why ADL may not improve when body functions and structures improve or why generalisation of skills does not readily occur (Reed et al., 2005). As a result of the limitations of the medical model, a biopsychosocial approach, which takes into account neuropathology, psychological responses, and inter- and intra-personal abilities, is recommended when working with people with TBI (Arlinghaus, Shoaib, & Price, 2005).

In 2001, the World Health Organization acknowledged the importance of a biopsychosocial perspective by endorsing the International Classification of Functioning, Disability, and Health (ICF; Peterson, 2005). The ICF, an alphanumeric coding system, was created to describe all health related states, irrespective of disability

(World Health Organization, 2001). It is an interactional model, with each component influencing and being influenced by the other components of the model (see Figure 1). In applying the ICF, the health condition of the person being assessed is first determined. A diagnosis can be made using the 10th revision of the International Classification of Disease (World Health Organization, 1992-1994) or the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (American Psychiatric Association, 2000). However, the diagnosis of the health condition is not critical in understanding the health-related functioning of the individual as the same health condition may present differently in different people or different health conditions may have the same presentation in different people (Bilbao et al., 2003). Nevertheless, it provides a starting point in assessing health-related functioning.

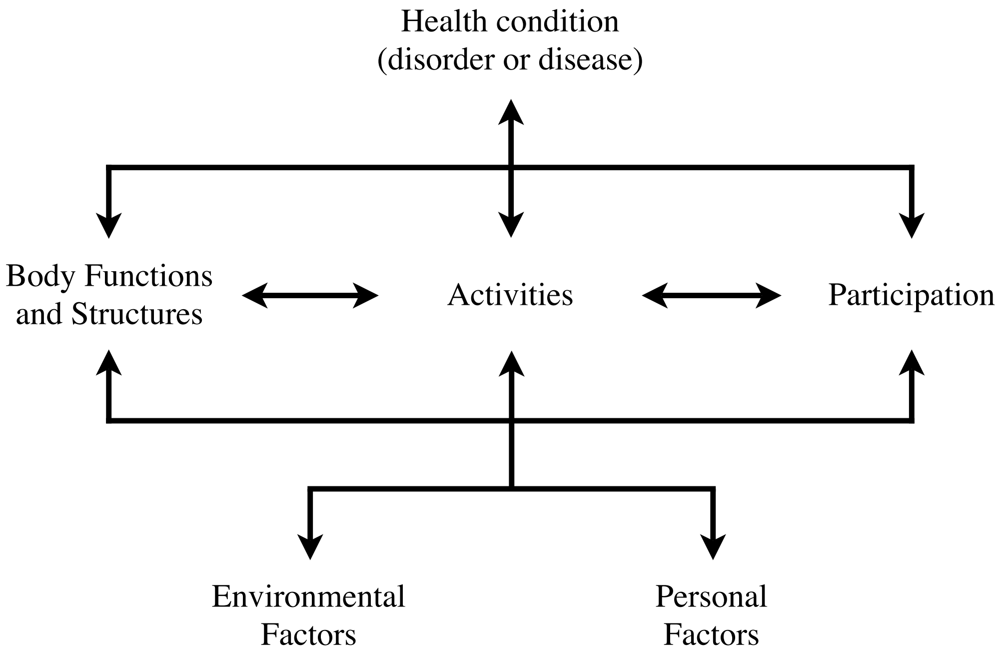

Figure 1. Interactions between components of the ICF (World Health Organization, 2001, p. 18).

The ICF is comprised of two main parts, with each part having two components. The first part, labelled Functioning and Disability, covers the topics Body Functions and Structure and Activities and Participation. Body Functions refer to both the physiological and psychological functions of the body, whereas Body Structures refer to physical structures that make up the human body. Each of these bodily aspects may be impaired to some degree. According to the ICF terminology, *impairments* are problems which cause a deviation from expected norms. Also subsumed under the label Functioning and Disability are Activities, which refers to the ability to complete a task, and Participation, which refers to being involved in societal contexts. These two terms are framed in positive language, with Activity Limitations referring to difficulties in completing tasks and Participation Restriction referring to difficulties experienced in societal contexts (World Health Organization, 2001).

Two qualifiers are needed when assessing Activity and Participation. These are Capacity, which refers to a person's ability to complete a task in a standardised environment, and Performance, which refers to a person's ability to complete a task in their present environment (World Health Organization, 2001). The distinction between these two qualifiers can be seen in people who are independent, despite displaying quite pronounced cognitive difficulties. For example, when tested in a standard environment, a person may fail to complete simple tasks adequately, but when compensatory strategies and assistive devices are used, the same individual may be able to complete quite complex tasks (Reed et al., 2005).

The second part of the ICF, labelled Contextual Factors, includes both environmental and personal factors. Environmental factors are all factors that are external to the individual, which includes their home environment and the societal environment they find themselves in, such as available rehabilitation services and social rules. Personal factors are idiosyncratic and can include attitudes and lifestyle (World Health Organization, 2001).

It is important to remember that the ICF was not designed as an assessment tool, rather it provides a framework to describe a person's functioning from various perspectives. It is left to the assessing clinician to determine what assessment measures will ultimately be used to create a formulation of the individual. However, incorporating the ICF in case formulation necessitates that the clinician consider functioning in both structured and natural environments, with and without compensatory interventions, and corroborate this information with significant other, or caregiver, feedback (Reed et al., 2005). With this information, interventions can then be designed to reduce impairments in body functions, activity limitations, participation restriction, and health-related barriers associated with environmental and personality factors (Bilbao et al., 2003).

Outcome Measures in Cognitive Rehabilitation

As discussed in a previous section, EF represents a constellation of functions that, when grouped together, represent higher-level cognitive functioning. Measures for assessing EF, such as the WCST and the Stroop Color-Word Test, have been used for decades (see Comalli et al., 1962; Milner, 1963) and are among the most commonly used measures in neuropsychological practice (Rabin, Burton, & Barr, 2007). The purpose of these, and other measures of EF, is to determine whether the cognitive processes involved in EF are impaired or not (Baron, 2004). This is done by comparing the results of a particular individual to a normative group. By definition, this form of assessment taps what the World Health Organization (2001) have termed impairment in body functions; psychological functions are being compared to a standard sample in order to determine whether those functions deviate from expected norms.

The purpose of cognitive rehabilitation following TBI is to provide people with the skills, or alter the environment sufficiently, to maximise independence in daily life. In other words, cognitive rehabilitation has the specific goal of reducing what the World Health Organization (2001) have termed activity limitations and participation restriction (Wilson, 2000). In current neuropsychological practice, there is a fundamental difference between the way cognitive functions are being assessed (body functions based) and how cognitive functions are being treated (activity and participation based; Odhuba, van den Broek, & Johns, 2005).

To bridge this difference, neuropsychologists have made an assumption about the ecological validity of formal neuropsychological tests. When someone performs poorly on particular measures of EF, it is assumed this individual will have difficulty in their daily life because the cognitive skills necessary in completing the measures of EF and the cognitive skills necessary in performing various daily activities are similar. The focus of cognitive rehabilitation is then to use retraining, compensatory strategies, and environmental manipulation to aid in the completion of these activities. Despite this assumption, research shows only modest correlations between formal measures of EF and everyday behaviours, with Pearson's correlations ranging from .24 to .50 (Burgess et al., 1998; Chaytor et al., 2006; Manchester et al., 2004).

There are several reasons for this modest correlation. First, many measures of EF were not designed with ecological validity in mind. Over the last several decades, referral questions posed to neuropsychologists have changed from being primarily diagnostic in nature to being asked to evaluate the functional ability of clients (Long, 1996; Rabin et al., 2007; Wilson, 1996). Although there has been a dramatic shift in the type of referral questions, the formal measures of EF used by neuropsychologists to answer these questions have remained the same. This is despite insufficient evidence for this practice (Burgess et al., 2006; Chaytor & Schmitter-Edgecombe, 2003).

Second, formal measures of EF are performed in standardised environments free from distraction and other confounding variables. This is done so that the best or optimal performance on a given measure can be achieved (Chaytor & Schmitter-Edgecombe, 2003). However, ecological validity refers to the association between a particular measure and everyday behaviour (Ready, Stierman, & Paulsen, 2001). Thus, ecological issues are concerned with what someone does rather than what they are capable of doing (McCue & Pramuka, 1998). A similar distinction between performance and capacity is made in the ICF, although these two qualifiers only apply to activities and participation and not to body functions or structures (World Health Organization, 2001). A person's cognitive capacity, or maximal functioning in a standardised environment, only tells us what they can do in that environment and may not allow neuropsychologists to predict that person's cognitive performance, or typical functioning in a day-to-day environment (Chaytor & Schmitter-Edgecombe).

Third, under formal test conditions, people are restricted from using the compensatory strategies taught in rehabilitation programmes. However, in situations outside the laboratory, these people are likely to use the skills they have been taught (Chaytor & Schmitter-Edgecombe, 2003). This will result in formal measures of EF underestimating cognitive performance (Chaytor et al., 2006; Long, 1996).

Associated with the third point is that formal measures of EF provide information on the cognitive inabilities people display rather than focusing on their cognitive abilities (Wilson, 1996). This focus on inabilities, or impairments in body functions, only provides information on what people are unable to do. This will likely underestimate cognitive performance if people display cognitive strengths, particularly if the cognitive strengths can compensate for cognitive difficulties (Chaytor & Schmitter-Edgecombe, 2003). For these reasons, formal measures of EF can only indirectly inform and assess rehabilitation outcomes (Wilson).

Finally, the standardised environment in which many formal measures of EF are conducted decrease the cognitive load on the very functions the measures were designed to assess (Holmes-Bernstein & Waber, 1990). Gioia, Isquith, and Kenealy (2008) explain that under formal testing situations, examiners often provide the structure and organisation to people and may even provide prompts so as to maximise an individual's performance. Thus, a person's score on formal measures may underestimate their cognitive performance because, in everyday situations, these people are required to plan and organise tasks on their own.

It should be noted that the issue of ecological validity of formal neuropsychological measures in some settings is moot. For example, obtaining measures of EF in research settings, regardless of the ecological validity of the tests used, satisfies the purposes of the research (J. M. Leathem, personal communication, March 22, 2007). Similarly, formal measures of EF will be an important component of any neuropsychological assessment undertaken to document impairment, such as in a psycholegal context (D. R. Babbage, personal communication, February 4, 2008). Therefore, it must be stressed that it is only when inferences are being made regarding a person's ability to perform everyday behaviours (e.g., during cognitive rehabilitation) that traditional measures of EF are inadequate.

Although traditional measures of EF are inappropriate for some uses in cognitive rehabilitation, there are some measures that specifically focus on people's performance in various everyday situations. Questionnaires, such as the Dysexecutive Questionnaire and the Behavior Rating Inventory of Executive Function (Gioia, Isquith, Guy, & Kenworthy, 2000) are easy ways to collect information about a wide range of everyday behaviours. Their use, however, may not provide an accurate depiction of this behaviour. This is particularly so in TBI populations as these people often display cognitive deficits, such as self-awareness difficulties, that prevent accurate reporting (Hart, Whyte, Kim, & Vaccaro, 2005; Turner-Stokes, 2002). Although proxy information is frequently obtained, this information is often different from client reports, and yet still may not be accurate (Tepper, Beatty, & DeJong, 1996). For example, Gioia and Isquith (2004) found only a moderate correlation (r = .32) between parent and teacher reports on the Behavior Rating Inventory of Executive Function—similar results have been found elsewhere in the literature (see Achenbauch, McConaughy, & Howell, 1987). Indeed, intra-familial distress, interpersonal relationships, awareness of the day-to-day functioning of the person with brain injury, date of injury, environmental influences, and stereotypes all affect information provided by others (Gioia et al., 2008; Manchester et al., 2004).

The Behavioural Assessment of Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie, & Evans, 1996) is an office-based test battery specifically designed to assess EF is real-life contexts. It is comprised of six tests and the Dysexecutive Questionnaire. Although the initial results found that the six BADS tests yielded significant correlations with the Dysexecutive Questionnaire, subsequent research has not found this to be the case (see Norris & Tate, 2000; Wood & Liossi, 2006). However, Norris and Tate did find that scores on the Action Programme Test, Zoo Map Test, and Modified Six-Elements Test could be used to predict performance on the Role Functioning Scale (McPheeters, 1984), a measure of activity and participation.

Taken together, these results suggest that the BADS is more ecologically valid that traditional neuropsychological measures.

Direct behavioural observation is another method of collecting information about people's ability to perform everyday activities. However, there are two main problems with using observation techniques in rehabilitation settings. First, they are time consuming (Turner-Stokes, 2002), and second, the functional ratings may be biased (Wolfson, Doctor, & Burns, 2000). This second point is particularly important in regards to the validity of observational assessments.

Bias in determining functional abilities can arise due to heuristics, or decision-making rules of thumb (Wolfson et al., 2000). Tversky and Kahneman (1974) explain that the representativeness heuristic is used when an assumption is made regarding the characteristics of Individual A, based on Individual A's similarity to Individual B. If Individual A is similar to Individual B, then it is assumed that they belong to the same group, whereas if they are dissimilar, it is assumed they belong to different groups. Predictions are then made about Individual A based on this perceived group membership. In terms of assessing functional outcomes, clinicians may rate a person's ability to be low if they know that the person has had a brain injury. This may be because executive dysfunction is common following TBI and, therefore, the person should display these impairments.

The availability heuristic occurs when, based on previous experience, an assumption is made regarding the probability of an action occurring (Tversky & Kahneman, 1974). In terms of assessing functional outcomes following TBI, if a rating clinician has seen people benefit from cognitive rehabilitation, any future clients that complete rehabilitation may also be rated with high ability by that clinician. This high rating can occur despite the client's actual ability (Wolfson et al., 2000).

Wolfson et al. (2000) conducted a study to determine the types of biases that occur when rating functional outcome using the Functional Independence Measure (FIM; Hamilton, Granger, Sherwin, Zielezny, & Tashman, 1987). They presented their participants, rehabilitation staff with FIM training, with four vignettes. Half of these vignettes contained potentially biasing information. Results suggest that when fictional team member FIM ratings were low, participants were biased to also give low FIM ratings. However, when fictional team member FIM ratings were high, participants provided accurate FIM ratings. Similarly, accurate FIM ratings were given when age and mental illness were used as potential biases. This study suggests that, under most circumstances, trained raters are able to provide accurate behavioural assessments.

Behavioural observation has been used for years in the assessment of Autism and Autistic Spectrum Disorders (see Lord et al., 1989). The Autism Diagnostic Observation Schedule - Generic (Lord et al., 2000) is one such measure that assesses an person's social and communicative behaviour using a variety of stimuli. This measure has been found to have high inter-class reliability, ranging from $r = .75$ to .97, and high inter-rater diagnostic concordance, ranging from 81% to 100% depending on the Module used. Furthermore, 95% of people with autism and 92% of people not on the autistic spectrum can be correctly classified by the Autism Diagnostic Observation Schedule - Generic. The relevance of this study in the current discussion shows that behavioural observation is a reliable and valid way of assessing cognitive ability.

Three commonly used assessment measures in rehabilitation settings in the United Kingdom, the Barthel index (Mahoney & Barthel, 1965), the FIM, and the FIM + Functional Assessment Measure (FAM; Hall, Hamilton, Gordon, & Zasler, 1993; Turner-Stokes, Nyein, Turner-Stokes, & Gatehouse, 1999), utilise direct observation techniques, with 83% of rehabilitation centres using at least one of these measures

(Skinner & Turner-Stokes, 2006). Of these measures, the Barthel index does not tap cognitive functioning (Turner-Stokes, 2002), the FIM does not adequately assess cognitive functioning (Malec, Moessner, Kragness, & Lezak, 2000), and the cognitive items in the FIM + FAM are subjective (Hall, Bushnik, Lakisic-Kazazic, Wright, & Cantagallo, 2001; Turner-Stokes, 2002) and do not assess the everyday behaviours that are often the focus of cognitive rehabilitation. In fact, there are very few measures that tap the cognitive aspects of everyday behaviours (Oddy, Alcott, Francis, Jenkins, & Fowlie, 1999), a situation that has not altered in subsequent years.

One such measure is a route-finding task devised by Boyd and Sautter (1993) to assess the ability of people with TBI to navigate in unfamiliar locations. Participants were asked to designate a place on a university campus that was unfamiliar to them and then navigate to that place. Two observers followed participants around the campus and rated each participant on task understanding, information seeking, retaining directions, error detection and correction, and on-task behaviour. Results found that inter-rater reliability was $r = .94$. Scores from the two observers correlated moderately with the Perceptual Organization ($r = .50$ and .49) and Verbal Comprehension ($r = .44$ and .53) indices of the WAIS-R. The results of this study suggest that direct observation can be used as a reliable way to assess EF and that measures of everyday functioning provide unique variance not accounted for by formal neuropsychological measures.

The Multiple Errands Test, Hospital Version (MET-HV; Knight, Alderman, & Burgess, 2002), is another measure of everyday EF in which people navigate in the "real-world". The traditional Multiple Errands Test (Shallice & Burgess, 1991) was modified in order to accommodate people who might find it difficult to navigate in public, such as people in wheelchairs or people with extreme behavioural difficulties. Participants in Knight et al.'s study were given a list of six tasks to perform within a 20 minute period. Scores on the MET-HV included the use of ineffective strategies in completing the various tasks, breaking task rules, implementation failures, task failures, and total errors. Performance was correlated with various formal neuropsychological measures. Inter-rater reliability of the MET-HV scores ranged from r = .81 to 1.00. Breaking task rules, task failures, and total errors correlated significantly with the perseverative errors score on the Modified WCST; r = .66, .49, and .67, respectively. No other significant correlations were found. Indeed, when corrections were made using the Bonferroni procedure, only breaking task rules and total errors remained significant. In terms of ecological validity, the MET-HV task failures score correlated significantly with Factor 2, Intentionality (r = .70), Factor 3, Executive memory (r = .53), and Total Score on the Ratings by Others scale of the Dysexecutive Questionnaire (r = .46).

Although route-finding and errand tasks have shown utility in assessing everyday behaviours (Boyd & Sautter, 1993; Knight et al., 2002), they may not be suitable for wide-spread use in cognitive rehabilitation centres. These procedures require time and resources that may not be available in smaller centres (Knight et al.). However, more importantly, these tests rely heavily on perceptual and spatial abilities. Thus, measures that remove extensive navigation may provide a more valid measure of EF. One such measure is the Test of Functional Executive Abilities (TOFEA; Bamdad, Ryan, & Warden, 2003).

The TOFEA scoring method is similar to Boyd and Sautter's (1993) route-finding task, in that participants are assessed on task understanding, information seeking, information retention, error detection and correction, on-task behaviour, and time management. In their study, Bamdad et al. (2003) instructed participants to gather information about a particular hotel. Some references were provided, such as a telephone, telephone directory, and maps, in order to accomplish the task. Moderate correlations were found between the TOFEA indices, with information retention and on-task behaviour having the lowest correlation with the other TOFEA indices. This was expected, however, given that task understanding, information seeking, error detection and correction, and time management tap EF, whereas information retention taps memory functioning and on-task behaviour taps attentional functions. The TOFEA yielded minimal, but significant, correlations to a variety of formal measures of EF suggesting that this measure may be useful in assessing everyday EF; Pearson's correlations ranged from $r = -.26$ to .18.

Chevignard et al. (2000) developed a functional outcome measure that took into account three different activities; shopping, cooking, and letter responding. In the first activity, participants were instructed that they were to buy the necessary ingredients to make scrambled eggs and a cake. They were given a recipe book, in order to plan what ingredients needed to be purchased, and sufficient money to purchase all of the ingredients. In the second activity, participants were instructed to make the scrambled eggs and cake from the ingredients that they had purchased. The final activity required participants to respond to a letter asking them to set-up an appointment. All of the necessary stationary was provided to participants, who were then required to post the letter. During each of these activities, two examiners directly observed each participant and assessed them based the components of EF as described by Lezak (1995). In addition to these scores, each participant's performance was reanalysed and scored based on context neglect, control errors, environmental adherence, distractibility, dependency, and behavioural disorder. The results found no significant correlation between performance on the daily tasks and formal neuropsychological tests or on self- and caregiver-reports on the Dysexecutive Questionnaire. However, a significant correlation was found between the daily tasks and the Dysexecutive Questionnaire when an occupational therapist completed the questionnaire. Despite the lack of significant correlations, participants with brain injury made significantly more errors on the everyday tasks than did healthy controls. This suggests that the task may potentially be measuring important variance in executive abilities that is related to everyday functioning and which is not captured by more traditional measurement techniques.

In two follow-up studies, Chevignard et al. (2008, 2009) again assessed EF by way of a cooking task. The procedure and scoring criteria for performance on these tasks were the same as used in Chevignard et al. (2000). In both studies, participants with TBI made significantly more errors than control participants. Although no significant correlations between the cooking tasks and formal neuropsychological measures were found in either study, Chevignard et al. (2008) found that nearly 50% of the variance of scores on the cooking task were accounted for by scores on the Six Elements task (26%), WCST (10%), Brown-Peterson Paradigm (6.7%; Van der Linden, Coyette, & Seron, 1992), and verbal fluency (6.5%; Benton, 1968).

In a similar study, Baguena et al. (2006) found that performance on a “real-life” task (cake baking) was significantly correlated with formal neuropsychological measures. Specifically, they found that errors on the daily task were positively correlated with Trail Making Test A (*r* = .74) and Trail Making Test B (*r* = .85). Furthermore, they also found that errors on the daily task were negatively correlated with performance on the Six Elements Test (*r* = -.78), sustained attention as measured by the KT Test (r = -.83; Barbizet & Duizabo, 1977), and the copy score on the Rey-Osterreith Complex Figure (*r* = -.84).

A major problem with the route-finding task (Boyd & Sautter, 1993), the MET-HV (Knight et al., 2002), the TOFEA (Bamdad et al., 2003), and the cooking tasks

(Baguena et al., 2006; Chevignard et al., 2000, 2008, 2009) are that they lack appropriate standardisation and rigorous testing. In contrast, the Assessment of Motor and Process Skills (AMPS; Fisher, 2001a, 2001b) is a measure of everyday behaviour that has been standardised with over 46,000 participants and has been in use by Occupational Therapists for over two decades. The AMPS was specifically designed to quantify the ability to perform everyday behaviours (Kottorp, Bernspång, & Fisher, 2003). People are observed performing an everyday activity, such as cooking or cleaning, and are scored based on two criteria. The first rates the person's motor skill (i.e., manipulating objects within the task), whereas the second rates the person's process skills on cognitive aspects of the task, such as planning, initiation, and problem solving (Fisher, 2001a). This measure provides information that many rehabilitation teams have found useful in understanding and quantifying a person's functional abilities, but falls short in terms of quantifying the fine-grained cognitive variables involved. The test designers explicitly state that the measure does not attempt to determine the underlying aspects of cognitive difficulties that may relate to the observed functional performance. Indeed, there is only one AMPS score for all of the cognitive functions that are used to complete a task - the "Process Skills" score. Someone may fail to complete a task because they have forgotten what the task was, were unable to cognitively shift to a different aspect of the task, or became distracted. Yet, this information is combined into a single process skills score. Thus, the formal scores generated from this measure cannot inform an analysis of the ways in which functional cognition is impaired, nor specifically assess the outcome of specific rehabilitation efforts designed to address them.

Other standardised measures similarly do not allow for a fine-grained analysis of cognitive ability. The Rabideau Kitchen Evaluation – Revised (Neistadt, 1992) assesses the ability to sequence the steps necessary in making a sandwich and hot drink. However, like the AMPS, this measure yields one score, ranging from 0 (independent) to 120 (assistance needed with every step), for all of the cognitive processes necessary for task completion (Josman & Birnboim, 2001). Although test-retest reliability of the Rabideau Kitchen Evaluation – Revised was $r = .80$ with 86% inter-rater agreement, the measure was only significantly correlated with Block Design from the WAIS-R ($r = -.60$), a measure of visuospatial and visuoperceptual skills.

In the Kitchen Task Assessment (Baum & Edwards, 1993) individuals are asked to make pudding on a stove. During performance, several components of EF are separately scored (i.e., initiation, organisation, step performance, sequencing, safety, and completion). Inter-rater reliability was found to be $r = .85$. Components of the Kitchen Task Assessment were significantly correlated to the Token Test Short Version ($r = -.31$ to -.68; De Renzi & Vignolo, 1962) and Trail Making Test A ($r = .39$ to .51). However, physical ability was required to complete the Kitchen Task Assessment. For example, if an individual needed physical assistance to complete parts of the task, they could not receive a score of independent. This was regardless of whether or not the individual displayed the cognitive ability to complete the task. Thus, scores on this measure are a combination of both cognitive and physical abilities. In addition, the Kitchen Task Assessment assesses a limited array of executive components (Josman & Birnboim, 2001) and was specifically designed for individuals with Alzheimer's dementia (Baum & Edwards)—it has not been used in TBI populations.

Finally, the Naturalistic Action Test (Schwartz, Segal, Veramonti, Ferraro, & Buxbaum, 2002) is another standardised measure in which a limited array of executive functions (i.e., sequencing and errors) are assessed during the performance of several everyday activities; preparing a simple meal, wrapping a gift, and organising a schoolbag. Inter-rater reliability, assessed using weighted kappa scores, ranged between .95 and 1.00. Naturalistic Action Test scores were significantly correlated with the cognitive subscale of the FIM for participants with right-hemisphere cerebrovascular accident, $r = .72$, and for participants with TBI, $r = .72$. Despite the limitations of this measure, promising research is underway which has extended the number of executive abilities measured by the Naturalistic Action Test in dementia populations to include error corrections (Bettcher, Giovannetti, & MacMullen, 2008).

Conclusion

People who experience a TBI will likely display a range of cognitive difficulties. Over the past several decades, neuropsychologists have developed numerous useful measures for assessing these cognitive difficulties (see Bamdad et al., 2003; Burgess, Alderman, Evans et al., 1996; Comalli et al., 1962; Hall et al., 1993). However, information gathered through formal neuropsychological assessment remains abstract, as the ecological validity of these measures is modest (Burgess et al., 1998; Manchester et al., 2004). Thus, there is a need to ground cognitive assessment in real-life behaviours. In terms of the ICF model (World Health Organization, 2001), an assessment based on activity limitation would benefit cognitive rehabilitation more than an assessment based on impairments in body functions, as cognitive rehabilitation is geared towards reducing activity limitations (Wilson, 1996).

The current study is an attempt to develop a systematic behavioural observation approach in quantifying ongoing cognitive activity limitations that tap EF after TBI. In order to accomplish this, the study was divided into three phases, each with different, but related, purposes. The aim of the first phase, the Pre-Study Survey, was to identify a common activity that a number of people with TBI appear to have difficulty performing.

Once a suitable activity was identified, a method needed to be developed that would provide a reliable way of assessing EF during the performance of said activity—the Developmental Trial. The final phase of the study, the Main Assessment, utilised the method that was created in the Developmental Trial with people who had sustained a TBI. The aim of the Main Assessment was to examine the performance of people with TBI on the cognitive activity limitations measure compared to performance on formal measures of cognitive impairments, self- and family member-reports (see Figure 2).

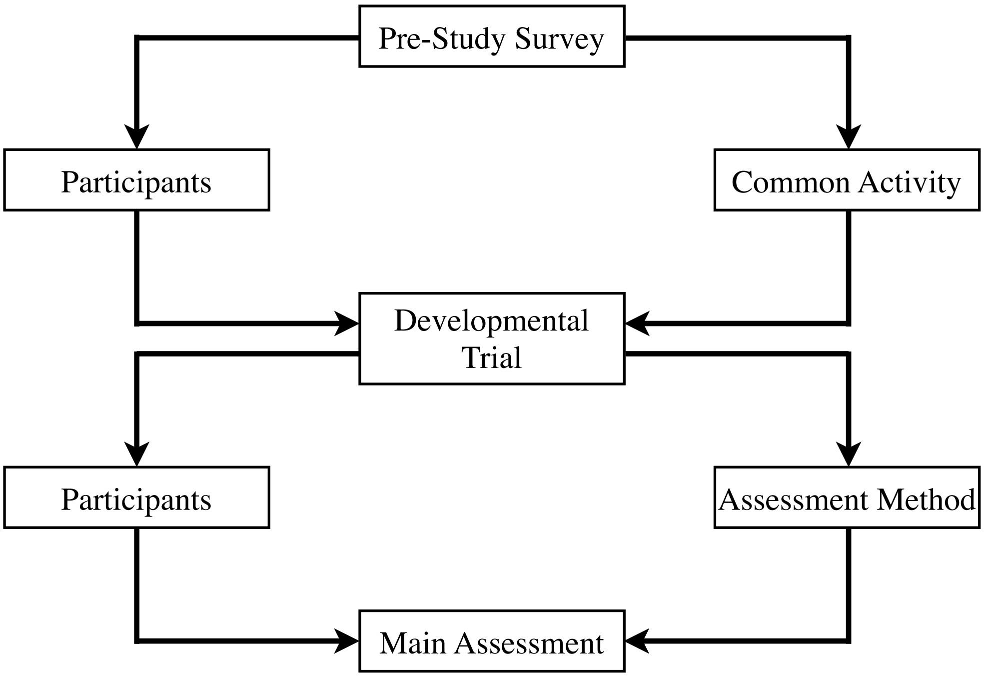

Figure 2. Relationship between the different phases of the study.

Chapter Two

Pre-Study Survey

As discussed in the previous chapter, there is a need to develop ecologically valid measures of EF. However, one issue facing the development of such measures is that ecological validity varies depending on who is being assessed and the everyday situations they find themselves in. Chaytor, Temkin, Machamer, and Dikmen (2007) explain that this is because ecological validity, or the association between a test and everyday behaviour, is dependent on the operational definition of everyday behaviour. For example, if everyday behaviour is defined as driving ability, then performance on a driving simulator task would likely have higher ecological validity than performance on a bed making task. On the other hand, if everyday behaviour is defined as instrumental activities of daily living, then performance on a bed making task may have higher ecological validity than performance on a driving simulator task. Thus, all measures, both traditional and those that are purported to be ecologically valid, are not actually ecologically valid in a general sense, but display increased ecological validity when the measures are associated with a specific everyday behaviour.

Before the Developmental Trial and Main Assessment could be conducted, a suitable "real-life" task needed to be identified. To this end, the Pre-Study Survey was developed to gather information regarding the types of daily activities performed by people with TBI in New Zealand. The purpose of the Pre-Study Survey was to find a task that could be considered an activity of daily life and was sufficiently complex to ensure that EF underlied a significant part of the successful completion of the task (see Stuss & Alexander, 2000). It was not intended that a formal analysis would be conducted with the data collected; rather, the use of a structured survey provided a systematic way of collecting information that was used to guide the development of the “real-life” task.

Method

Participants

The inclusion criteria for potential participants were: (a) they had received a TBI after the age of 16; (b) had completed their acute medical rehabilitation following their brain injury and their treatment team considered they were no longer in post-traumatic amnesia; (c) had satisfactory English language skills, assessed qualitatively in interview; and (d) had satisfactory visual acuity, assessed by participant self-report. Family members were included in the Pre-Study Survey as it was anticipated that the participants would be involved in the Developmental Trial or Main Assessment. Inclusion criteria for family members included being someone who knew a participant sufficiently to comment on their everyday behaviours. Participants were not excluded if they did not have a suitable family member or support person.

Thirteen people (7 male, 6 female) with TBI participated in the Pre-Study Survey. Ethnicity was determined by participant self-identification; ten were New Zealand European, one was Māori, and two were European. Twelve participants were employed at the time they sustained their brain injury; seven were employed at the time of testing. Participants were on average 43 years old at testing and on average eight years post injury (see Table 2).

Table 2

Demographic Characteristics of Participants in the Pre-Study Survey

Variables		
Mean age at test in years (SD)	42.7	(12.8)
Mean age at injury in years (SD)	35.0	(14.4)
Mean time since injury in months (SD)	93.2	(90.6)
Sex		
Male (%)	7	(54)
Female (%)	6	(46)
Ethnicity		
New Zealand European (%)	10	(77)
Other European (%)	2	(15)
Māori (%)	1	(8)
Employment		
Employed at time of injury (%)	12	(92)
Employed at time of test (%)	7	(54)

Participants were recruited from past and current clients of the Massey University Psychology Clinic at Wellington and from the Brain Injury Association of New Zealand. Current clients of the Massey Psychology Clinic were contacted by their treating clinician, whereas past clients were referred by the Clinic Director to the researcher, who then contacted potential participants. An article was also published in *Brainlink* (Lewis & Babbage, 2007a), the Brain Injury Association of New Zealand's newsletter, giving a brief synopsis of the research and asking for people interested in participating in the project to contact the researcher.

Ten past or current clients of the Massey Psychology Clinic were contacted but declined to participate, one had moved, and four contact phone numbers had been disconnected. One person contacted the researcher after reading the *Brainlink* article, but did not meet inclusion criteria. Final participants were thus twelve past or current clients of the Massey Psychology Clinic and one participant self-referred from the *Brainlink* article.

Although injury severity data was not collected in the Pre-Study Survey (see Limitations of the Pre-Study Survey, p. 48), an estimate of severity can be made. The 12 participants recruited from the Massey University Psychology Clinic had likely sustained a severe TBI, whereas the severity of injury for the remaining participant who was self-referred remains unknown.

Measures

Activities of daily life questionnaire. Seventeen questions, 16 closed- and one open-ended, were used to determine commonly performed daily activities. The questions were derived based on a review of the European Brain Injury Questionnaire (Svendsen, Teasdale, & Pinner, 2004; Teasdale et al., 1997), the AMPS (Fisher, 2001a, 2001b), and from Olver, Ponsford, and Curran (1996), and covered activities that in the opinion of the researcher could be recreated in a laboratory setting. In addition, an open-ended question was used to enquire about any rehabilitation strategies used to help complete the daily activities. A response sheet, with six response choices, was printed in black 18-point Times Roman font on a white A4 sized sheet of paper (see Appendix A, p. 140). It was anticipated that participants with TBI may have difficulty remembering the six response choices. Thus, to aid in the administration of the questionnaire, a response sheet was given to participants.

European Brain Injury Questionnaire (EBIQ). The EBIQ (Svendsen et al., 2004; Teasdale et al., 1997) is a self- and proxy-report questionnaire used to assess physical, emotional, and cognitive functioning. Nine hundred and five people with brain injuries (571 with cerebrovascular accident, 258 with TBI, and 76 with other brain injuries) and their relatives and 203 controls and their relatives participated in Teasdale et al.'s normative study. People with brain injuries were recruited from six European countries and Brazil. They had a mean age of 47.6 (standard deviation of 17.8) years, were 62% male, had a mean education level of 4.6 (on a Likert-type scale where 1 = illiterate and 7 = obtained a university degree), and were on average 31.8 months post injury. Control participants who were recruited from France and Brazil, had a mean age of 41.8 (standard deviation of 11.9) years, were 48% male, and had a mean education level of 5.4.

Teasdale et al. (1997) found that the Cronbach's Coefficient Alpha for the nine EBIQ subscales ranged from .47 to .90 for self-report ratings from people with brain injuries and from .54 to .92 for their relatives' ratings. They considered that these results indicated acceptable internal consistency, as all but two subscales (self-reported Isolation and Communication) yielded Cronbach's Alphas above .5. However, Bland and Altman (1997) suggest that Alphas of .7 are necessary to be considered satisfactory. When stricter criteria are used to assess internal consistency, only four subscales, Cognitive, Impulsivity, Depression, and Core, yielded Alphas above .7 for both self-report and relative's rating. The remaining subscales yielded relatively low Cronbach's Alphas indicating that the items within the subscales displayed relatively low internal consistency or may not be assessing a unidimensional construct. Despite this, Teasdale et al. also found that the EBIQ could be used to discriminate between people with brain injuries and controls (eight subscales yielded a $p < .01$), between dysphasia from non-dysphasia (the communication subscale yielded a $p < .01$), between cerebrovascular accident and TBI (five subscales yielded a $p < .05$), and time since injury (six subscales yielding a $p < .05$). Furthermore, the EBIQ is frequently used to assess outcomes following cognitive rehabilitation (B. A. Wilson, personal communication, February 25, 2008).

In the original EBIQ, examinees are asked to mark one of three circles that correspond to the three answers. The layout of the questionnaire was modified for this study to reduce reliance on working memory by writing the three available answers in full following each question. This was intended to increased the ease of completing the questionnaire and to minimise recording mistakes. A response sheet, with the three response choices (Not At All, A Little, and A Lot), was printed in black 55-point Times Roman font on a white A4 sheet of paper (see Appendix B, p. 146).

Demographic Questionnaire. A brief self-report demographics questionnaire was used to assess age, date of the TBI, ethnicity, and several other demographic characteristics (Appendix C, p. 150). The questionnaire content was based on the New Zealand government's 2006 general population census.

Procedure

Informed consent was obtained before testing began. (See Appendix D, p. 151, and Appendix E, p. 159, for the information sheets and consent forms, respectively.) Participants completed the Activities of Daily Life Questionnaire, EBIQ, and demographic questionnaire. Participants were given a response sheet and asked to rate their current ability to perform the activities to be read by the examiner. The Activities of Daily Life Questionnaire was then verbally presented to the participants. It was anticipated that participants with TBI may have difficulty completing the questionnaire independently due to impairments such as working memory deficits. Thus, the questionnaire was read to participants to ensure ease of administration and completion.

Following each question, participants were asked to note how often the activity occurred during the past week. After all of the questions had been answered, the 16 activities and any additional activities were reviewed to determine the strategies used to help complete the activities.

Participants were told that the EBIQ covered a range of problems that people sometimes experience. They were given a response sheet and asked to respond to each question by indicating how often they experienced the problem during the past month. The questionnaire was then presented verbally by the examiner. The same rationale for verbally presenting the Activities of Daily Life Questionnaire was used for verbally presenting the EBIQ. Although this differs from the standard administration of the EBIQ, in which participants complete the questionnaire independently (Teasdale et al., 1997), it was also thought that verbally presenting the questionnaire would reduce any fatigue effects that may result in completing the EBIQ independently. One participant was posted the information sheet, consent form, and questionnaires and asked to complete the study independently. They were informed that if they had any questions, they could contact the researcher who would endeavour to answer them. Following the EBIQ, participants completed the demographic questionnaire, which was verbally presented to participants.

Family members either completed the Activities of Daily Life Questionnaire and EBIQ in the presence of the researcher or on their own. Family members who met with the researcher were given the Activities of Daily Life Questionnaire and asked to rate their family member's ability to complete the activities as well as to indicate how often the activity occurred during the past week. They then completed the Activities of Daily Life Questionnaire independently. Following the completion of this questionnaire, family members completed the EBIQ independently, as per the instructions in Svendsen et al. (2004) and Teasdale et al. (1997). On one occasion, the Activities of Daily Life Questionnaire and EBIQ were presented verbally for a family member, along the lines of the administration for participants. This was done as a couple worked together to complete the family member versions of the questionnaires. In another case, one participant contacted a family member and asked them to complete the questionnaires.

Planned Analysis

A comprehensive analysis of the data was not intended. Rather, the information gathered was used to help guide the development of a functional, "real-life", task that would be appropriate for use in the subsequent study. Thus, it was anticipated that descriptive statistics of the dependent variables would be suitable for this purpose. The dependent variables from the EBIQ were the *Somatic*, *Cognitive*, *Motivation*, *Impulsivity*, *Depression*, *Isolation*, *Physical*, *Communication*, and *Core* subscales. The dependent variables from the Activities of Daily Life Questionnaire were the ability scores and frequency ratings for each of the 16 items.

Results

Participant scores ranged between 1.58 (Physical subscale) and 2.24 (Cognitive subscale), where a score of 1.00 indicated that the problem or difficulty was not experienced over the last month and a score of 3.00 indicated that the problem or difficulty was experienced a lot over the last month. Family member scores ranged between 1.67 (Depression subscale) and 2.15 (Cognitive subscale). There were no significant differences between participants and their family members on any of the EBIQ subscales (see Table 3).

Participants reported that they were independent at completing the Activities of Daily Life Questionnaire items *Make a Cup of Coffee or Tea* (*M* = 1.00), *Post a Letter at the Post Office* (*M* = 1.00), *Go to the Bank* (*M* = 1.00), and *Separate the Recycling* (*M* = 1.00), whereas they reported that they were the least independent for the item *Make a Complex Meal that Requires Cooking* (*M* = 2.17). Family members reported that participants were the most independent for the item *Post a Letter at the Post Office* (*M* = 1.13), whereas they reported that participants were the least independent at *Managing Your Finances* (*M* = 2.56), followed by *Make a Complex Meal that Requires Cooking* (*M* = 2.50; see Table 4). Significant differences were found between participants and their family members' ratings on the participant's ability to go to the bank, use the yellow or white pages, and separate recycling. No other significant differences were found between participant and family member ratings (see Table 4).

Participants reported that the item *Use Public Transportation* occurred the fewest number of times over the past seven days (*M* = .18), whereas *Make a Cup of Coffee or Tea* occurred the most (*M* = 3.38). Similarly, family members reported that the item *Use Public Transportation* occurred the fewest number of times (*M* = .25) and *Make a Cup of Coffee or Tea* occurred the most (*M* = 3.14). No significant differences were found between participants and family members on the number of times that each item on the Activities of Daily Life Questionnaire had been completed over the previous seven days (see Table 5).

The types of compensatory strategies used to help complete the items on the Activities of Daily Life Questionnaire varied between participants. Ten participants reported that they used at least one compensatory strategy to *Make a Complex Meal That Requires Cooking* and to *Pay the Bills*, whereas two participants reported that they used at least one compensatory strategy to *Make a Cup of Coffee or Tea*, *Go to the Bank*, and *Use Public Transportation* (see Table 6).

Table 3

Comparisons Between Participants and Family Members on the EBIQ Subscales

	TBI Group (*n* = 13)		Family Member Group (*n* = 9)				
Variables	M	SD	M	SD	t	p	d
Somatic	2.00	.47	1.99	.40	.12	.90	.05
Cognitive	2.24	.35	2.15	.45	.49	.63	.21
Motivation	1.74	.21	1.69	.51	.32	.75	.14
Impulsivity	1.88	.22	2.10	.37	-1.79	.09	.78
Depression	1.81	.45	1.67	.54	.68	.50	.30
Isolation	1.98	.31	1.97	.59	.04	.97	.02
Physical	1.58	.48	1.83	.55	-1.16	.26	.50
Communication	1.96	.57	1.89	.64	.28	.78	.12
Core	1.91	.28	1.93	.44	-.14	.89	.06

Table 4

Comparison Between Ability Scores on the Activities of Daily Life Questionnaire as Reported by Participants and Family Members

	TBI Group (*n* = 13)		Family Member Group (*n* = 9)				
Variables	M	SD	M	SD	t	p	d
Complex meal	2.17	1.64	2.50	1.69	-.44	.67	.20
Manage your finances	1.91	1.58	2.56	1.51	-.93	.67	.42
Organise daily routine	1.85	.80	1.89	.93	-.12	.91	.05
Drive a vehicle	1.67	1.56	2.33	2.00	-.86	.40	.38
Pay the bills	1.54	.97	2.38	1.60	-1.50	.15	.68
Do Housework	1.54	.88	2.33	1.41	-1.63	.12	.71
Go grocery shopping	1.50	1.17	1.88	.83	-.78	.44	.36
Simple meal	1.46	.97	1.67	1.32	-.42	.68	.18
Use Public transportation	1.33	.82	2.00	1.41	-1.02	.33	.56
Use the internet	1.30	.67	2.22	1.48	-1.78	.09	.82
Hygiene	1.23	.44	1.33	.71	-.42	.68	.18
Phonebook	1.08	.29	1.78	.83	-2.70	.01	1.19
Go to the bank	1.00	.45	1.67	.87	-2.23	.04	1.00
Post letter	1.00	.00	1.13	.35	-1.13	.28	.53
Separate recycling	1.00	.00	2.00	1.41	-2.25	.04	1.07
Coffee or Tea	1.00	.00	1.22	.67	-1.22	.24	.53

Note. Complex meal = make a complex meal that requires cooking; Simple meal = make a simple meal that requires no cooking; Hygiene = maintain personal hygiene; Phonebook = use the yellow or white pages to look up a phone number; Post letter = post a letter a the post office; Coffee or Tea = make a cup of coffee or tea. The higher mean score, the less independence in completing the activity.

Table 5

Comparison Between Occurrence of Activities of Daily Life Questionnaire Items as Reported by Participants and Family Members

	TBI Group (*n* = 13)		Family Member Group (*n* = 9)				
Variables	M	SD	M	SD	t	p	d
Coffee or Tea	3.38	.87	3.14	1.46	.47	.65	.22
Hygiene	3.25	.62	3.50	.84	-.72	.48	.36
Drive a vehicle	3.20	1.23	2.00	1.67	1.66	.12	.85
Simple meal	2.54	1.13	2.00	1.41	.93	.36	.44
Do Housework	1.92	1.19	.83	.75	2.05	.06	1.01
Use the internet	1.91	1.38	2.33	1.37	-.61	.55	.31
Organise daily routine	1.83	1.03	1.80	1.30	.06	.96	.03
Pay the bills	1.45	.82	1.00	1.00	.96	.35	.52
Complex meal	1.42	1.38	1.20	1.30	.30	.77	.16
Separate recycling	1.17	1.03	1.20	.84	-.06	.95	.03
Go grocery shopping	1.00	.63	1.20	.45	-.63	.54	.34
Phonebook	.75	.45	.83	.41	-.38	.71	.19
Manage your finances	.73	.65	1.17	.75	-1.27	.23	.64
Post letter	.64	.67	.83	.75	-.55	.59	.28
Go to the bank	.60	.97	.83	.41	-.56	.59	.29
Use Public transportation	.18	.60	.25	.50	-.20	.84	.12

Note. see Table 4.

Table 6

Number of Participants Who Used Compensatory Strategies for Activities of Daily Life Questionnaire Items (n = 13)

Variables	Number of Participants Using Strategies	Examples of strategies used
Make a complex meal that requires cooking	10 (77%)	“Recipes”, “Keep things simple”, “Use microwave”
Pay the bills	10 (77%)	“Attach money to the bill with paper a clip”, “Automatic payments”
Manage your finances	9 (69%)	“Use the internet”, “Allowed $100 per week for my wallet”
Organise daily routine	9 (69%)	“Write everything down”, “Use a diary and whiteboard”
Go grocery shopping	9 (69%)	“Think ahead”, “Use lists”
Separate recycling	7 (54%)	“Specific bags for specific items”, “Use recycling box”
Make a simple meal that requires no cooking	6 (46%)	“Use notes”, “Husband helps”
Do Housework	5 (38%)	“Cleaning lady”, “Vacuum part of the house and do the rest latter”
Maintain personal hygiene	5 (38%)	“Parents remind me”, “Same routine everyday”
Use the internet	5 (38%)	“Get help from son”, “Going to computer course”
Use the yellow or white pages to look up a phone number	4 (31%)	“Read slowly”, “Parents look up number if I can’t find it”
Drive a vehicle	4 (31%)	“Drive in the morning or after a rest”, “Avoid rush hour”
Post letter	3 (23%)	“Use the same places”, “Use positive thoughts”
Make a cup of coffee or tea	2 (15%)	“Put sugar and tea in cup first, then add water”
Go to the bank	2 (15%)	“Use the internet”, “Have credit cards”
Use Public transportation	2 (15%)	“Use only on off-peak times”

Discussion

The purpose of the Pre-Study Survey was to gather information that would inform the development of a "real-life" task to be used in the subsequent studies. Both participants and a nominated family member completed the survey. Participant reports did not significantly differ from family member reports on any of the subscales of the EBIQ or on 12 of the 16 items on the Activities of Daily Life Questionnaire. This suggests that the participant group did not display the self-awareness difficulties that are typically reported following TBI (see Flashman, Amador, & McAllister, 2005), at least in the areas covered by the survey. However, Svendsen et al. (2004) found that participants and family members only differed significantly on the Motivation subscale of the EBIQ and researchers have found that accurate self-awareness after TBI is dependent on the skills being assessed (Prigatano, Altman, & O'Brien, 1990), time since the injury, and severity of the brain injury (Dirette & Plaisier, 2007).

Pagulayan, Temkin, Machamer, and Dikmen (2007) found that participants with TBI displayed more self-awareness deficits on the Psychosocial Factor of the Sickness Impact Profile (Bergner, Bobbitt, Pollard, Martin, & Gilson, 1976) than on the Physical Factor of the scale. Similarly, Prigatano et al. (1990) found that participants displayed more accurate self-awareness for activities of daily life than for psychosocial aspects of functioning. Although the EBIQ does assess aspects of cognitive, psychological, and social functioning, the Activities of Daily Life Questionnaire only assesses activities of daily life. Therefore, self-awareness deficits should have been reduced in the current sample due to the specific types of questions contained within the questionnaires. Pagulayan et al. (2007) also found that 46% of participants displayed self-awareness deficits at one month post-injury, whereas 33% of participants displayed self-awareness deficits at one-year post-injury. The mean time since injury of participants in the Pre-

Study Survey was 93 months. Thus, any self-awareness deficits that the current participants may have experienced immediately following their brain injuries would have decreased by the time they were tested. Finally, the severity of injury was not assessed in the Pre-Study Survey and, thus, it is unknown whether or not the majority of participants sustained mild TBIs, as mild TBI is associated with more accurate self-awareness than moderate or severe TBI (Dirette & Plaisier, 2007).

The scores on the EBIQ indicate that participants experience moderate problems or difficulties in each of the subscales. Both participants and their family members reported the most difficulty in the cognitive domain. This supports McCullagh and Feinstein's (2005) claim that cognitive impairments are the predominant feature following TBI.

Participants indicated that they were the least independent in making a complex meal that required cooking, whereas family members indicated that participants were least independent in managing their finances. Ponsford, Harrington, Olver, and Roper (2006) found that after two years post-injury participants with TBI continued to required assistance in managing their finances and in completing cooking tasks. Both participants and family members indicated that cooking a complex meal and managing finances were each completed several times per week. Thus, these two tasks represent everyday tasks that are performed by people with TBI in New Zealand.

Stuss and Alexander (2000) explain that performing complex and novel tasks, as well as adapting learned information to new tasks, requires the use of EF. Ten participants indicated that they used rehabilitation strategies to complete a complex cooking task, whereas nine participants indicated that they used rehabilitation strategies to help them manage their finances. This suggests that these tasks are sufficiently complex to ensure that EF underlied a significant part of the successful completion of the tasks.

After reviewing the information gathered from the Pre-Study Survey, a complex cooking task was chosen as the “real-life” task to be used in the subsequent studies. There were several reasons for making this choice. First, it was assumed that participants would enjoy the novelty of a cooking task and, therefore, be more likely to participate in the research than if the “real-life” task was a financial management task. Second, it was considered unlikely that participants would have prepared the selected cooking task. This ensured that preparing the dish would be novel, thus ensuring that EF would underlie a significant part of the successful completion of the task. And third, several studies that have attempted to develop ecologically valid measures of EF also used cooking tasks (see Baguena et al., 2006; Chevignard et al., 2000, 2008, 2009).

Limitations of the Pre-Study Survey

One of the limitations of the Pre-Study Survey was the small number of participants and family members who took part. However, this is mediated by the fact that the study was intended as a means of gathering information to ensure that the “real-life” task that was created for the subsequent studies had face validity and was a suitable task.

A second limitation was that data regarding injury severity was not collected in the current series of studies. This occurred due to an oversight on the part of the researcher, rather than a purposeful attempt to not collect or report this information. However, an estimate of injury severity can be made. Those participants who were past or current clients of the Massey University Psychology Clinic at Wellington likely sustained a severe TBI. This is because clients entered the service by way of a contract that the clinic had with the Accident Compensation Corporation, a state-owned accident compensation agency, which stated that people needed to have sustained a severe TBI before being initially referred. Similarly, clients recruited from the residential rehabilitation service likely also sustained severe TBIs due to having a similar arrangement with the Accident Compensation Corporation.

Chapter Three

Developmental Trial

The purpose of this phase of the study was to develop a systematic behavioural observation approach for quantifying ongoing cognitive activity limitations that tap EF after TBI. Information gathered from the Pre-Study Survey helped identify a cooking task as an appropriate functional "real-life" activity to be used in the Developmental Trial. Thus, the aim of the Developmental Trial was to develop a reliable method of quantifying cognitive activity limitations, using systematic behavioural observation of a cooking task. This was done by creating a scoring method prior to any participants being seen and then revising of the method, by way of trial and error, following each participant's performance. It was not intended that the validity of the method would be assessed during the trial.

Ethical Considerations

Standard ethical principles, and in particular the Code of Ethics for Psychologists Working in Aotearoa/New Zealand (Code of Ethics Review Group, 2002), guided this research. There were several specific ethical principles considered during the current series of studies. First, there were concerns for the physical safety of participants, particularly given that the functional task was a cooking task. Therefore, a task needed to be carefully chosen to ensure participant safety was maintained. More generally, participant confidentiality was to be maintained, attempts were made to make the testing environment as comfortable for participants as possible, and support persons could be present during the assessment sessions. Associated with this second point was ensuring that participants received information regarding the outcome of the study. Therefore, participants were to indicate whether or not they wanted feedback on the study—the content of this feedback related to the overall results of the study, rather than to individual performance during the study. Fourth, the possibility that emergency situations may arise was anticipated. Therefore, an experienced Clinical Psychologist from the Massey University Psychology Clinic at Wellington was available for consultation. Finally, ethical approval was sought from a state-accredited ethics committee.

Method

Participants

The inclusion criteria for potential participants were: (a) they had sustained a TBI after the age of 16; (b) had completed their acute medical rehabilitation following their brain injury and their treatment team considered they were no longer in post-traumatic amnesia; (c) had satisfactory English language skills, assessed qualitatively in interview; and (d) had satisfactory visual acuity, assessed by participant self-report. Inclusion criteria for family members included being someone who knew a participant sufficiently well to comment on their everyday behaviours. Participants were not excluded if they did not have a suitable family member or support person.

Ten people (six male, four female) with TBI participated in the Developmental Trial. Of these participants, seven indicated that they were New Zealand European, one indicated that they were Māori, and two indicated that they were European. Seven participants were employed at the time they sustained their brain injury, with four being employed at the time of testing (see Table 7). All of the participants in the Pre-Study Survey were invited to take part in the Developmental Trial. Three participants accepted the invitation and took part in the Developmental Trial.

Table 7

Demographic Characteristics of Participants in the Developmental Trial

Variables	
Mean age at test in years (SD)	46.40 (14.49)
Mean age at injury in years (SD)	36.13 (12.64)
Mean time since injury in months (SD)	70.00 (57.34)
Sex	
Male (%)	6 (60)
Female (%)	4 (40)
Ethnicity	
New Zealand European (%)	7 (70)
Other European (%)	2 (20)
Māori (%)	1 (10)
Employment	
Employed at time of injury (%)	7 (70)
Employed at time of test (%)	4 (40)

Participants were recruited from past and current clients of the Massey University Psychology Clinic at Wellington, the Brain Injury Association of New Zealand, and a residential rehabilitation service. Current clients of the Massey Psychology Clinic were contacted by their treating clinician, whereas past clients were referred by the Clinic Director to the researcher, who then contacted potential participants. All clients from the residential rehabilitation service were initially contacted by staff from that service. If the client agreed to participate, contact details were then passed to the researcher. An article was published in *Brainlink* (Lewis & Babbage, 2007a)—the same article that was described in the Pre-Study Survey. After the conclusion of the Pre-Study Survey, all participants who contacted the researcher after having read the article in *Brainlink* were invited to take part in the Developmental Trial. A second article was published in the Brain Injury Association of Wellington's newsletter (Lewis & Babbage, 2007b), giving a brief synopsis of the research and inviting people interested in participating in the project to contact the researcher. Seven participants were past or current clients of the Massey Psychology Clinic, two were referred from a residential rehabilitation service, and one participant self-referred from the *Brainlink* article. Although injury severity data was not collected it is likely that nine participants likely sustained a severe TBI, with the severity information for the self-referred participant being unknown.

Measures

The Activities of Daily Life Questionnaire, EBIQ, and Demographic Questionnaire were described in the Pre-Study Survey (see Chapter Two, p. 33)

Formal Neuropsychological Measures. The Sorting Test, Alternate Form, and the Tower Test from the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001) were used as formal measures of EF. These stand-alone tests are suitable for assessing EF in people with traumatic brain injuries (Homack, Lee, & Riccio, 2005) and were included in the Developmental Trial as it was anticipated that after a reliable method of quantifying cognitive activity limitations had been developed that the validity of the method would be assessed. Therefore, although these measures were administered in the Developmental Trial, the scores obtained on the measures are analysed in the Main Assessment (see Chapter Four, p. 75 and p. 76, for a more in-depth description of the Sorting Test, Alternate Form, and the Tower Test, respectively).

Functional Task. Microwave brownie making was chosen as part of the cooking task. A chocolate brownie recipe (Downs, n.d.) was modified to reduce the overall cost and to simplify the task (see Appendix F, p. 165). More specifically, an ingredient that was deemed optional was eliminated and the equipment required was marginally simplified. This recipe was selected for a number of reasons. First, preparing chocolate brownies mirrored previous research in which participants were asked to make a cake (coincidentally also chocolate; Baguena et al., 2006; Chevignard et al., 2000, 2008, 2009). Second, two participants surveyed indicated that they cooked almost exclusively with a microwave oven. Therefore, the use of a microwave increased, in some cases, the ecological validity of the task. Associated with this second point was avoiding creating a task that might place participants in a dangerous situation (e.g., being burned by a hot oven element). This was a particular concern given that the cognitive difficulties experienced by people following a TBI can impair their ability to safely perform everyday tasks (see Landa-Gonzalez, 2001). The use of a microwave and microwavable dish decreased this hazard as neither the microwave nor the microwavable dish would become hot during cooking.

Fourth, the portability of a microwave cooking task allowed the researcher to meet participants at a location convenient to them (e.g., at home, at a residential rehabilitation service, or at the Massey Psychology Clinic), transporting a microwave oven itself if necessary. One of the initial goals of the project was to create a scenario in which people with TBI could be observed performing everyday behaviours for a relatively short period of time. The microwave brownie task suited this goal as the brownies could be prepared in less than 30 minutes. Thus, the brevity of this task has the potential to be readily incorporated as part of an assessment that also involves traditional psychometric assessment measures. Finally, it was anticipated that cooking a chocolate treat might make more sense to participants than asking them to cook another standard meal in-between their regular meals, and might be an enjoyable experience for them.

The second part of the cooking task was making a hot drink. This task was chosen as each participant surveyed indicated that they had made at least one cup of coffee or tea within the last seven days. Adding a second part to the cooking task required participants to shift from the first aspect, chocolate brownie making, to a new aspect—making a hot drink does not automatically follow from the brownie recipe and requires participants to shift their cognitive set. Participants also needed to monitor their performance of the cooking task to ensure that they made the hot drink at the appropriate time. It was also deemed useful to include an everyday task that could be performed without the need for a recipe; Chevignard et al. (2000, 2008) provided a chocolate cake recipe but also required participants to prepare scrambled eggs without the aid of a recipe, for instance. The hot drink making task could be adapted by the participant, who could either make the tea that was provided or they could make another hot drink (i.e., one using ingredients they provided if testing was being conducted in their homes or with ingredients that the Psychology Clinic routinely provided if testing was being conducted at the Psychology Clinic). There is obvious face validity to the request that participants prepare a hot drink alongside baking chocolate brownies.

Two manuals were created. The first was an administration manual outlining the administration of the functional task, the second a scoring manual used to evaluate participant performance based on nine aspects of EF that were derived from Boone et al. (1998), Burgess et al. (1998), Busch et al. (2005), Lezak et al. (2004), and Sohlberg and Mateer (2001). During the development process, there were five instances where additional information was added to the scoring manual to clarify the original scoring criteria. Furthermore, a number of changes were also made—one to the administration manual and 17 to the rating scale (see Appendix G, p. 166, and Appendix H, p. 169, for the changes to the Administration Manual and Scoring Manual, respectively).

The one change to the administration manual occurred after the first participant had completed the task. It was intended that at some point during the brownie making task that participants would make the hot drink. In this way participants would have to monitor their performance to ensure that they switched between the two tasks (i.e., from the brownie making task to the hot drink making task and then back to the brownie making task). However, the first participant made the hot drink before making the brownie. Therefore, from the second participant onward, participants were instructed to make the hot drink while the brownies were cooking. This also took advantage of the natural break in the brownie making task when participants were waiting for the brownie to cook.

Eleven of the 17 changes to the scoring manual related to the method of assigning a functional rating after participant performance had been scored. Briefly, the initial scoring manual did not specify the exact number of cognitive activity limitations that needed to occur in order for a specific functional rating to be awarded. However, it was deemed necessary, in order to create a reliable measure, for each functional rating to be associated with a specific set of criteria—the number of cognitive activity limitations performed (see Appendix I, p. 189 for a detailed description of the revision process).

Planning and organisation involved identifying, organising, and sequencing the steps or elements of a task sufficiently for the task to be completed successfully (see Lezak et al., 2004; Sohlberg & Mateer, 2001). Scoring this aspect was essentially a measure of sequencing as identifying and organising refer to internal, non-observable, processes. However, it was assumed that if participants correctly sequenced the steps of the task, they had also correctly identified and organised those steps.

Task steps were arranged based on the order they occurred in the recipe. A step was considered correct if it was: (a) the next uncompleted step in the sequence; or (b) any previously uncompleted step, as this allowed participants to adjust their planning and organisation of the task based on their current situation (see Appendix I, p. 189, paragraph 1). However, there were four sections within the task where the sequencing, as set-out in the recipe, did not have to be followed in order for a participant to achieve a correct sequencing score. These sections were created as correctly sequencing these steps did not affect the outcome of the task. For example, although the recipe stated that white sugar should be added before brown sugar, the order in which participants added sugar did not affect the outcome as participants were to stir the mixture after adding the sugars (see Appendix I, p. 189, paragraph 2).

A planning and organisation error was recorded if an additional step (i.e., a step not included in the recipe) detracted from the successful completion of the task. For example, stirring the mixture after adding white sugar, but before adding brown sugar, is an additional step. However, as this step does not affect the outcome of the task, it was thought that it should not be considered an error. This is in contrast to adding a second cup of flour, which would cause the brownie to become too dry and negatively affect the outcome (see Appendix I, p. 190, paragraph 1).

Executive memory related to task outcome and the ability of people to complete each step or element of the task correctly (see Busch et al., 2005). A step was considered correct if it was completed as specified in the recipe. Additional steps that detracted from the successful completion of the task were scored as errors. For example, adding chocolate chips to the brownie mixture would enhance the outcome of the brownies and would not be scored as an error, whereas adding canola oil would negatively affect the outcome of the brownies and, therefore, would be scored as an error (see Appendix I, p. 190, paragraph 2).

Initiation involved starting goal-directed behaviour, either through self-initiation or examiner prompting, such that an individual's plan was translated into action (Busch et al., 2005; Sohlberg & Mateer, 2001; Appendix I, p. 191, paragraph 1). Cognitive shifting referred to the ability to move from one element of the task to another element of the task (see Boone et al., 1998; Lezak et al., 2004; Appendix I, p. 191, paragraph 2). Impulsivity involved performing impulsive behaviours (Appendix I, p. 191, paragraph 3). Directed and sustained attention referred to the ability to maintain one's attention on the task, despite external or internal distractions (see Burgess et al., 1998; Sohlberg & Mateer). The current operational definition did not differentiate between external or internal distractions, only that the participant's focus had been directed away from the task (see Appendix I, p. 191, paragraph 4).

Error detection involved monitoring one's performance such that errors became apparent (see Burgess et al., 1998; Busch et al., 2005; Lezak et al., 2004). Error detection scores were allocated based on the number of errors performed (see Appendix I, p. 192, paragraph 1). Four error detection criteria were recorded, which included monitoring (i.e., checking to ensure that a behaviour was correct), false alarms (i.e., claiming that an error had occurred when in fact they did not perform an error), errors detected (demonstrated through either error correction itself or explicit statements that indicated the participant was aware an error had occurred), and failure to detect errors (see Appendix I, p. 192, paragraph 2). Error correction involved correcting detected errors and was scored based on the number of error detections that occurred (see Appendix I, p. 192, paragraph 3). Finally, time management involved spending appropriate amounts of time on various aspects of the task such that the task was completed in the allotted time (see Boone et al., 1998; Lezak et al., 2004; Appendix I, p. 193, paragraph 1).

Participants were given a functional rating based on their performance on the Developmental Trial. Ratings included *Full Functioning, Functioning with Some Limitations*, *Achieved Some Functional Outcomes*, *Significant Disability*, and *Apparent Total Disability*. A rating of *Full Functioning* was awarded for each component of EF if participants displayed no cognitive activity limitations. The remaining functional ratings for six of the nine components of EF (planning and organisation, executive memory, initiation, cognitive shifting, impulsivity, and directed and sustained attention) were assigned to one of four functional ratings based on the number of examples of cognitive activity limitations displayed (see Figure 3; Appendix I, p. 193, paragraph 2).

The cut-off points for the functional ratings were initially selected somewhat arbitrarily, as there was no external criterion available to help make the decision. The rationale was that there were 21 steps or elements of the task. It was assumed that if participants performed these steps without displaying executive difficulties, their performance suggested intact executive skills (i.e., full functioning). The remaining 21 steps were separated roughly into quarters, with three 5 steps sections and a six step section, which were preliminarily associated with a functional rating (see Appendix I, p. 193). As stated, these cut-off points were essentially arbitrary but held reasonable face validity for the researchers and provided at least a starting point for the coding of responses.

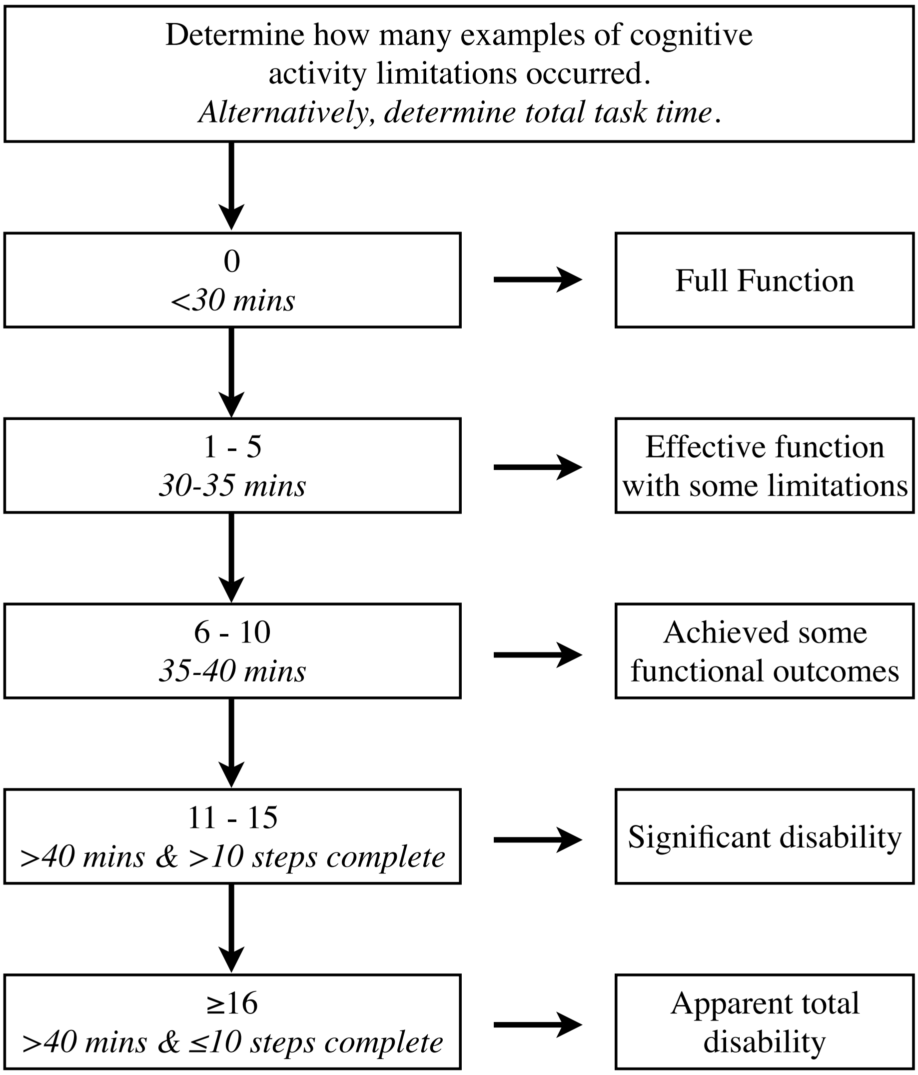

Figure 3. Decision tree for awarding functional ratings for planning and organisation, executive memory, initiation, cognitive shifting, impulsivity, directed and sustained attention, and time management.

The main error detection score was the percentage of performed errors that were detected. Similarly, the main error correction score was the percentage of detected errors that were corrected. Percentage values were given to these ratings, rather than simply recording actual detections and corrections, in order to provide a comparison between detected and undetected errors and between corrected and uncorrected errors.

Two rating scales were created. The first was used if participants made three or fewer errors or if participants made three or fewer corrections, whereas the second rating scale was used if participants made more than three errors or made more than three corrections during their performance of the task (see Figure 4). Dividing the rating systems between three and four errors or corrections occurred as it was thought that performing three or fewer errors or corrections would not provide the examiner with enough data to conclude that a participant's functional ability was *Apparent Total Disability*. However, if a participant performed four or more errors or corrections, it was concluded that this functional rating may be appropriate (see Appendix I, p. 194, paragraph 1).

The cut-off points for the functional ratings that were awarded if participants performed four or more errors or corrections were similar to that of the six components of EF described earlier (i.e., functional ratings were assigned to one of five groups, with the same rationale used to create the percentage groupings). However, if participants performed three or fewer errors or corrections, functional ratings were assigned to one of four groups (see Figure 4). The rationale for creating these percentage groups was that if *Apparent Total Disability* was inappropriate, participants would receive a functional rating of *Significant Disability* if no errors were detected or corrected. The remaining functional ratings were associated with a percentage value (see Appendix I, p. 194, paragraph 2).

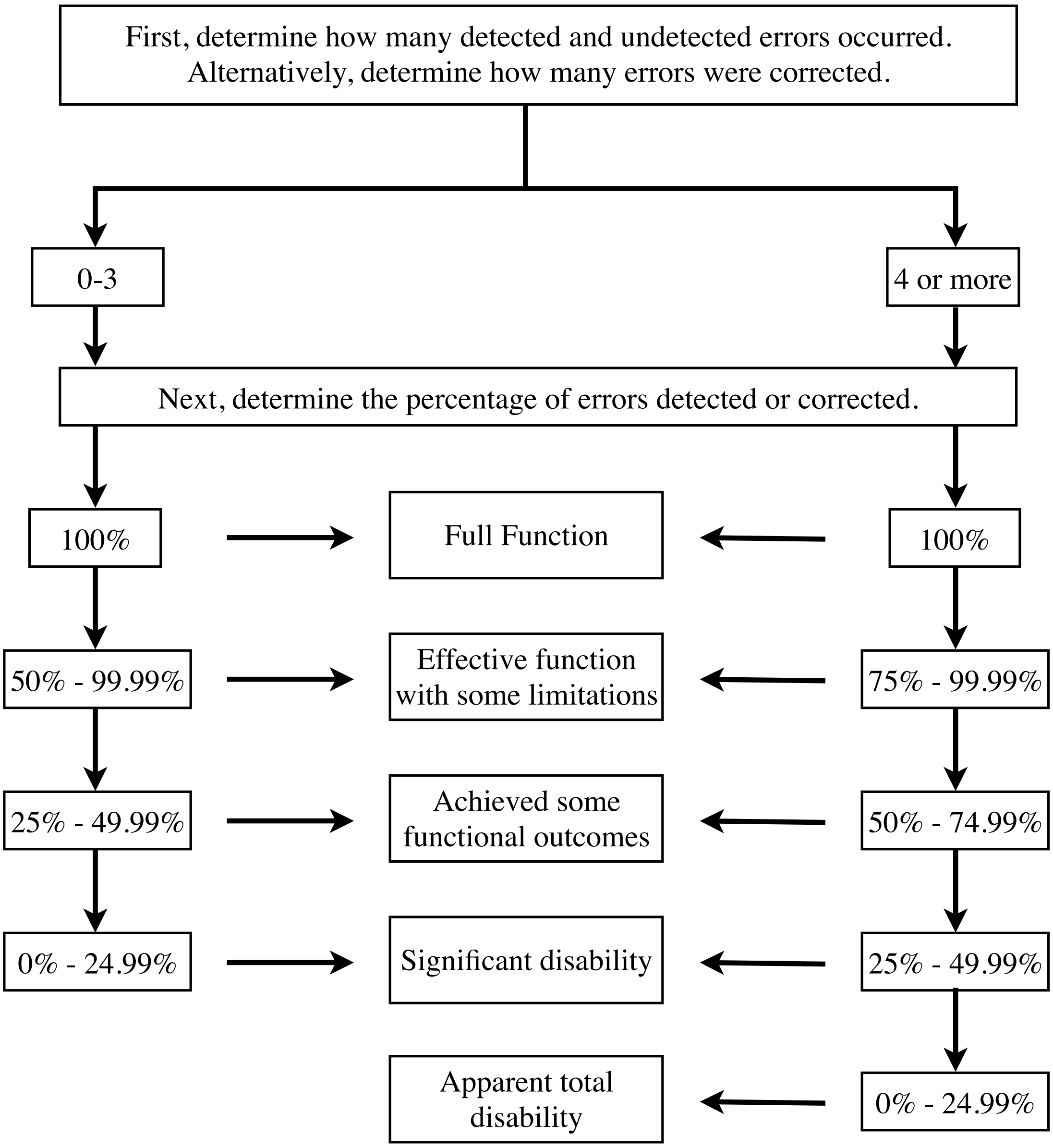

Figure 4. Decision tree for awarding functional ratings for error detection and error correction.

Participants were awarded a functional rating of *Full Function* for time-management if they completed the task within 30 minutes; Downs (n.d.) stated that their task could be completed within 30 minutes and the present researcher was able to completed the current task in less than 18 minutes. Additional time was given over the 30 minute mark in order to allow people to complete the task. If participants had not completed the task within 40 minutes, they were instructed to stop the task. As shown in Figure 3, when this occurred, the number of steps completed was used determine the functional rating (see Appendix I, p. 195, paragraphs 1 & 2).

Participants were given a subjective rating for each of the aspects of EF based on the quality of their behaviours in supporting the functional rating (termed, the *Evidence Score*). This rating was based on the following scale: (a) unequivocal; (b) strong evidence; (c) moderate evidence; (d) limited evidence; and (e) no evidence. The purpose of the evidence score was to provide the examiner with the ability to rate whether or not they felt that there was behavioural evidence to justify the functional rating. For example, if a participant required three examiner attention prompts, they would be given a functional rating of *Effective Function With Some Limitations* for the Sustained and Directed Attention aspect of EF. If the examiner believed that this rating was justified, they would record that there was strong evidence in the functional rating. Evidence scores were recorded for the first four participants before being abandoned due to the ambiguity of scoring (see Appendix I, p. 196, paragraphs 1 & 2).

A final subjective rating (termed, the *Confidence Score*) was used to assess the examiner's confidence in the functional rating. The confidence ratings were *very strong confidence*, *strong confidence*, *moderate confidence*, *limited confidence*, and *no confidence in prediction*. The purpose of this score was to allow the examiner to compare the functional rating with their subjective assessment of participant performance and then to determine whether or not the two assessments matched. Confidence scores were also only recorded for the first four participants (see Appendix I, p. 197, paragraph 1).

A 5-point Likert-type scale was used to assess each participant's baking experience and confidence in baking, as well as to determine how well they thought they performed the task and their satisfaction with the outcome. Examiners rated the taste of the brownie (see Appendix J, p. 198). (Brownie was transported to the other examiners within 24 hours of the task completion and stored in an airtight container until consumption for scoring.)

Procedure

Pilot Trial. The rating scale was initially trialled with a healthy individual who feigned executive dysfunction. This person was coached to feign executive dysfunction, in order to provide the examiners with a very basic trial for data collection; it was not intended to provide a useful analogue to any genuine results. This person's performance on an everyday task was rated and then reviewed with her to determine whether or not the executive deficits she was mimicking were identified. The review revealed that the functional rating accurately matched her level of simulated executive dysfunction.

During this initial trial, the healthy individual was monitored performing a task by one examiner and was video-recorded in order for the second examiner to assess her performance. Based on the results of this initial trial, an additional step was implemented to increase inter-rater reliability. The Developmental Trial required participants to manoeuvre in a kitchen. The problem arose that in order to clearly see what participants were doing (e.g., which utensils they were using and the exact amount of each ingredient) the camera needed to be focussed quite closely on them. As a result, some behaviours took place 'off camera'. Therefore, behaviours and situations that the live rater thought might need clarification were noted and given to the video rater prior to their viewing of the tape. For example, the live rater would note cooking times in order to help the video rater score the appropriate steps correctly. (See example form recording this information for one participant, reproduced in Appendix K, p. 199).

A further issue that became clear as a result of the initial trial was that it would be difficult, at least in the early stages of the research, for examiners to accurately score live participant performance. It was anticipated that some performance details would be missed while examiners were recording previous target behaviours. Furthermore, it was not feasible to have two examiners at each live performance and yet it was imperative that the inter-rater reliability of the behavioural observation method be assessed. Filming participants for subsequent assessment is a commonly used procedure in studies that employ behavioural observation in order to counter these and similar issues (see Dath et al., 2004; Fricke, Unsworth, & Worrell, 1993; Hermansson, Bodin, & Eliasson, 2006; Lindén, Boschian, Eker, Schalén, & Nordström, 2005). Of note, Eranti, Janakiramaiah, Gangadhar, and Subbakrishna (1998) found that despite reductions in scores by the examiners who viewed the video-tapes, their scores correlated significantly with those of the live raters. This led them to conclude that video recording performance for later scoring is a valid and reliable method of assessing participant behaviours. Based on these findings, filming participants for later assessment was employed in the current study.

Developmental Trial. This study was administered in the context of a larger investigation into developing a method of quantifying cognitive activity limitations through systematic behavioural observation. Informed consent was obtained before testing began (see Appendix L, p. 200, for the information sheets and consent forms). Participants were asked to designate a family member for participation in the study. The functional task was conducted first, followed by the questionnaires from the Pre-Study Survey, the Sorting Test, and the Tower Test. The study had approval from a state-accredited ethics committee. One formal memory measure was also administered to these participants, as part of data collection for a separate (though related) study that will be reported elsewhere.

All of the ingredients and utensils needed to complete the cooking tasks were placed on a counter-top. Participants were asked to prepare some brownies, with the aid of a recipe supplied by the examiner. They were instructed to start by making the brownie and make the hot drink while the brownies were cooking. Participants were further informed that they were allowed to use any strategies they currently used to complete tasks of this nature. They were asked to work as quickly and as efficiently as possible. Following the completion of the task, participants rated their baking experience and their perceived performance on the task. Both the examiner and the participant rated the taste of the brownie.

Information regarding a participant's planning and organisation, executive memory, initiation, and cognitive shifting abilities were recorded following the completion of each step or element of the task. Information regarding impulsivity, directed and sustained attention, error detection, and error correction abilities were recorded only when specific behaviours were displayed. Time-management was recorded at the end of the task.

The procedure for the Activities of Daily Life Questionnaire, EBIQ, and Demographic Questionnaire was described in Chapter Two (p. 33) and the Sorting and Tower Tests were conducted as per the instructions in Delis et al. (2001). Participants were given NZ$20 in partial recognition of their travel expenses, time, and effort in participating in the research project. Typically, testing was conducted across two sessions lasting approximately one hour each, with these session separated across several days or weeks.

Two raters scored participant performance. The examiner scored participant performance both during the live performance of the cooking task and during the subsequent video replay—the second viewing took place to ensure that all relevant participant behaviours had been recorded. The second rater scored participant performance from the videotaped footage. Before viewing participant performance, the second rater was given the 'off camera' notes. Throughout the video replays, both raters were allowed to stop the video at any time or replay any section of the video in order to clarify participant performance; this was however only required for a small percentage of the footage viewed.

Each component was given a functional, evidence, and confidence rating. After independently scoring participant performance, the rating scale was evaluated to determine whether or not modifications were necessary. If modifications were deemed necessary, changes were made and implemented with subsequent participants.

Planned Analysis

The inter-rater reliability of the scoring system was assessed by way of intraclass correlation coefficient (ICC). This coefficient was chosen instead of Pearson's product-moment correlation coefficient (*r*) as the latter correlation coefficient does not take into account actual scores, but rather it is the relative proportion between the participants and their given scores that determines the size of the correlation coefficient (see Sheskin, 2004). For example, if the examiner rated the first four participants' planning and organisation scores as being 1, 1, 2, 1 (1 = Effective Function and 2 = Effective Function with Some Limitations) and the second rater gave these participants the scores 3, 3, 4, 3 (3 = Achieved Some Functional Outcomes and 4 = Significant Disability), the resulting Pearson's *r* would be 1.00. This indicates a perfect relationship between the raters and suggests that the measure is reliable. However, the first participant received a score of 1 from the examiner and a score of 3 from the second rater, the second participant received a score of 1 from the examiner and a score of 3 from the second rater, etc. Thus, the rater's scores for each participant are actually different, which suggest that there is, in fact, a poor relationship between the raters. This is reflected in the resulting ICC (.11).

The ICC model used in the current study was a two-way random effects model, using absolute scores (see Appendix M, p. 213). This model was chosen, from the six possible ICC models, as raters independently rated each participant's performance. In addition, the two raters in the Developmental Trial were chosen from a sample of convenience, with the assumption that their scores would be similar to hypothetical other raters. That is, it was assumed that if different raters scored participant performance, their rating would be similar to the scores given by the raters in the Developmental Trial. Absolute scores were chosen instead of the mean scores as the actual ratings were of interest. ICCs were interpreted according to Haneline (2007), in which scores less than .40 were considered poor, scores between .40 and .75 were considered fair to good, and scores above .75 were considered excellent. (See Shrout & Fleiss, 1979, for an in-depth discussion regarding the six ICC models).

Despite the advantage of ICCs over Pearson's correlations, ICCs will be low when scores are restricted in range. This will occur even if there is high actual rater agreement (Graham & Naglieri, 2003). For this reason, percentage agreements were used to assess rater agreement for those variables that yielded low ICCs. Percentage agreements were calculated based on a procedure described by Birkimer and Brown (1979; see Equation 1).

$$\left(\frac{\text{number of agreements}}{\text{number of agreements } + \text{ number of disagreements}}\right) \text{ x } 100 \qquad (1)$$

The dependent variables from the functional task were the nine functional ratings, namely *Planning and Organisation*, *Executive Memory*, *Initiation*, *Cognitive Shifting*, *Impulsivity*, *Sustained and Directed Attention*, *Error Detection*, *Error Correction*, and *Time Management*.

Results and Discussion

The aim of the Developmental Trial was to develop a reliable method of quantifying cognitive activity limitations, using systematic behavioural observation of a cooking task. It should be stressed that the validity of the method was not assessed in the Developmental Trial. The process used in the Developmental Trial mirrored the developmental process of other assessment measures (see DeVellis, 2003). In particular, the scoring criteria was: (a) initially informed by a literature review; (b) piloted with people with TBI; and (c) evaluated to ensure that inappropriate aspects of the scoring criteria were discarded and that appropriate aspects were either introduced or retained.

The scoring manual was refined after each participant's performance and, therefore, the results reflected the average agreement between examiners throughout the revision process. According to Haneline's (2007) evaluation criteria, the ICCs for the Planning and Organisation (1.00), Time Management (1.00), Error Detection (.95), and Impulsivity (.76) were within the excellent range, whereas the ICCs for Sustained and Directed Attention (.55) and Executive Memory (.54) were within the fair to good range. The ICCs for the remaining three functional ratings, Cognitive Shifting (.23), Error Correction (.04), and Initiation (.00), were within the poor range.

Percentage agreements were calculated for three functional ratings in order to determine whether or not a reduced range of scores led to the ICC scores being in the poor range. Percentage agreement between raters for Initiation, Error Correction, and Cognitive Shifting were 90%, 80%, and 70%, respectively. These scores suggest that a reduction in range produced the low ICCs, rather than poor inter-rater agreement.

The results of the Developmental Trial provided an initial indication that the behavioural observation method of quantifying cognitive activity limitation was reliable. Therefore, based on this evidence, a larger scale assessment of the reliability of the final version of the rating scale was warranted. In addition, an examination of the validity of the method was deemed appropriate.

Limitations of the Developmental Trial

One of the limitations of the Developmental Trial was the small number of participants that took part. This has several important implications. There is an increased probability that the participant group does not represent the general population of people with TBI. Given that the scoring manual was adjusted based on participant performance, the relatively few participants in the Developmental Trial may have skewed the assessment method such that it may not be applicable to the general population of people with TBI. However, the small sample size of the Developmental Trial is mitigated by the fact that the purpose of the trial was simply to develop a reliable measure, with the validity of the method being assessed in the Main Assessment (see Chapter Four, p. 72). Furthermore, small pilot studies are routinely used to modify assessment methods prior to larger scale evaluations (see DeVellis, 2003; Spence, Moss-Morris, & Chalder, 2005).

The small sample size also increased the probability of chance agreements between the two raters. This is of particular importance as percentage agreements were used to assess inter-rater reliability for three of the ability scores. Furthermore, the use of percentage agreements as a measure of inter-rater reliability is problematic in that this statistic does not take into account chance agreements between raters. Each participant was capable of receiving one of five functional ratings. If one assumes that there is a 20% probability that raters will agree simply by chance, as participants were given one of five possible functional ratings, agreements of between 70% and 90% would indicate good inter-rater reliability. However, no participant received a functional rating of Apparent Total Disability and very few received ratings of Significant Disability and Achieved Some Functional Outcomes. This suggests that the distribution of possible scores were not uniformly spread, which has the effect of increasing the probability of chance agreements (see Hayes & Hatch, 1999).

Chapter Four

Main Assessment

The Developmental Trial found that the newly developed method of quantifying cognitive activity limitations through systematic behavioural observations displayed satisfactory inter-rater reliability. Based on these results, an analysis of the validity of the method was undertaken. Thus, the Main Assessment had the following two main aims. First, to compare the performance of people with TBI on the functional measure to their performance on both formal measures of cognitive impairments as well as to self- and family member-reports. Second, to compare the performance of people with TBI to people without brain injuries on the functional "real-life" measure. In addition to these aims, a research assistant was trained in the scoring procedure in order to determine the ease with which a new rater could be trained to provide reliable scores.

Method

Participants

The inclusion criteria for potential participants with TBI and their family members and the method of recruitment were reported in the Developmental Trial (see Chapter Three, p. 50). In addition, a new article was published in *Brainlink* (Researchers launch, 2009, April) giving a brief synopsis of the research and asking for people interested in participating in the project to contact the researcher.

Twenty seven people (18 male, 9 female) with TBI participated in the Main Assessment. Of these participants, 21 indicated that they were New Zealand European, two indicated that they were English, one indicated that she was both Māori and European, one indicated that he was Chinese, and one indicated that she was an American. Twenty participants were employed at the time they sustained their brain injury, with fourteen being employed at the time of testing (see Table 8). Although injury severity data was not collected in the Main Assessment, 12 participants were recruited from the Massey University Psychology Clinic, 10 were recruited from the residential rehabilitation service, four self-referred from Brain Injury Association of New Zealand, and one self-referred after reading about the study on the Massey University website. Thus, 22 participants likely had a severe TBI, whereas severity information for the remaining five participants remains unknown.

Nine participants from the Developmental Trial participated in the Main Assessment—their performance on the functional task was re-evaluated using the final version of the rating scale. Two participants completed the functional task but left the inpatient rehabilitation unit before the formal neuropsychological measures were administered; one of these participants also did not complete the questionnaire portion of the study and, therefore, demographic information regarding this participant is unknown.

Inclusion criteria for potential participants in the normative group were that they: (a) were over the age of 16; (b) had never had a TBI; (c) had satisfactory English language skills (assessed informally in interview); and (d) had satisfactory visual acuity (assessed by participant self-report). Sixteen people (8 male, 8 female) participated in the Main Assessment (see Table 8). There were no significant differences between the TBI and normative groups on any of the demographic characteristics.

Participants from the normative group were recruited through a press release from Massey University (Lewis & Gibson, 2009). A feature story was published in The Dominion Post, a national daily newspaper, describing the research project and the journalist's experience of participating in the study as a control subject (Newton, 2009, March 16). Finally, each participant was asked to invite three friends or family members to participate in the study.

Table 8

Demographic Characteristics of Participants in the Main Assessment.

Variables	TBI Group (*n* = 26)	Normative Group (*n* = 16)
Mean age at test in years (SD)	49.1 (18.5)	48.5 (15.7)
Mean age at injury in years (SD)	41.8 (18.0)	
Mean time since injury in months (SD)	70.5 (89.3)	
Sex*		
Male (%)	18 (67)	8 (50)
Female (%)	9 (33)	8 (50)
Ethnicity		
New Zealand European (%)	21 (81)	6 (37)
Māori (%)		3 (19)
Other European (%)	2 (7)	3 (19)
Other Ethnicity (%)	3 (12)	4 (25)
Employment		
Employed at time of injury (%)	20 (77)	
Employed at time of test (%)	14 (54)	9 (56)
Education		
No formal qualification (%)	2 (8)	1 (6)
Secondary degree (%)	6 (23)	4 (25)
Tertiary degree (%)	11 (42)	10 (63)
Trade certificate (%)	7 (27)	1 (6)

Note. *n = 27 for the TBI group.

Measures

The Activities of Daily Life Questionnaire, EBIQ, and demographic questionnaire were described in the Pre-Study Survey (see Chapter Two, p. 33).

Sorting Test. The Sorting Test, Alternate Form, from the D-KEFS (Delis et al., 2001) was used as a formal measure of EF. This stand-alone test is suitable for assessing EF in people with traumatic brain injuries (Homack et al., 2005). The Alternate Form was used rather than the Standard Form as participants with TBI were more likely to have completed the Standard Form during a formal neuropsychological assessment following their TBI. Therefore, the Alternate Form was selected to reduce the likelihood of practice effects.

The Sorting Test, Alternate Form, was standardised with 286 people from the United States of America. Participants had a mean age of 47.5 years and a standard deviation of 23.5 years. Although specific reliability analyses for the Alternate Forms are not provided in their *Technical Manual*, Delis et al. (2001) found that correlations between the Alternate Form and the Standard Form ranged between $r = .39$ and .72. The Standard Form was standardised with 1,750 people and was stratified based on the United States of America government's 2000 general population census. Internal consistency was assessed by splitting the test into two parts and then using the Spearman-Brown formula to correct the Pearson's correlation coefficients. The resulting coefficients ranged between .55 and .86 and had a standard error of measurement ranging between 1.13 and 2.01. Test-retest reliability was found to range between $r = .43$ and .73 (Delis et al.).

Validity data for the Sorting Test, Alternate Form, is not provided by Delis et al. (2001) in their *Technical Manual*, rather all of the data on the validity of the Sorting Test pertains to the Standard Form of the measure. However, Delis et al. used Angoff's

(1984) design-IIB linear calibration method to obtain scores on the Alternate Form that were equivalent to scores on the Standard Form. Normative information was then developed and Delis et al. state that the calibration procedure produced an Alternate Form that is directly comparable to the Standard Form. Research using the Sorting Test, Standard Form, found that the primary measures from the Sorting Test yield low to moderate relationship with the WCST. In addition, performance on the Sorting Test can differentiate between people with Alzheimer's and Huntington's disease (Delis et al.). The Sorting Test provides a measure of problem-solving, initiation, inhibition, and conceptualisation ability and was conducted as per the instructions in Delis et al.

Tower Test. The Tower Test from the D-KEFS (Delis et al., 2001) was used as a formal measure of EF. This stand-alone test is suitable for assessing EF in traumatically brain injured populations (Homack et al., 2005). Internal consistency for the Tower Test was calculated the same way internal consistency was calculated for the Sorting Test (i.e., using the Spearman-Brown formula to correct Pearson's correlation coefficients) and ranged from .43 to .84 and had a standard error of measurement ranging from 1.21 to 2.27. Test-retest was found to range between $r = .38$ and .51 (Delis et al.).

In their pilot study, Delis et al. (2001) found that the Tower Test was able to differentiate between Alzheimer's and Huntington's disease. In addition, the primary measure (total achievement score) displayed low to moderate correlations with the WCST. The Tower Test provides a measure of planning and maintenance ability, as well as the ability to inhibit impulsive responses. This test was conducted as per the instructions in Delis et al.

Functional Task. The cooking tasks, administration manual, and scoring manual were described in Chapter Three (p. 50). In addition, a subjective analysis was introduced in the scoring manual for the Main Assessment. Participants were given a rating of unequivocal, strong, moderate, limited, or none based on their ability to perform each of the nine aspects of EF. For example, if a participant did not display any Planning and Organisation errors, their performance suggests that they showed unequivocal capacity to perform the task and at the same time showing no (none) incapacity on the task. On the other hand, if a participant displayed three Executive Memory errors, their performance may suggest that they showed strong capacity and moderate incapacity on the task (see Appendix N, p. 214, and Appendix O, p. 231, for the final version of the scoring manual and record form, respectively).

A research assistant was trained to score participant performance. Training took place over two sessions, lasting two hours each. During a third session, the research assistant independently scored a participant's performance in the presence of the examiner. This final session was conducted to ensure that the assistant was fully trained in the scoring procedure.

In the first training session, the scoring manual was reviewed with the research assistant as she had not read the scoring manual prior to the session. Specific examples of typical cognitive activity limitations for each of the nine aspects of EF were described to the assistant (e.g., if a participant adds too much flour to the mixture, but removes some of the flour before mixing the ingredients together, they have performed an error detection). This took approximately one hour. The remainder of the session involved reviewing a participant's performance via videotaped footage. During this initial video review, the research assistant was shown specific examples of cognitive activity limitations, how to record these on the record form, and how the behaviours were examples of a cognitive activity limitation as described in the scoring manual. For example, immediately after the participant had measured the margarine, the videotape was paused and the impulsive behaviour was highlighted—the participant used one of his own coffee cups to measure the margarine despite being told that all utensils had been provided. The scoring manual was then reviewed to show how the participant's behaviour was similar to the examples of impulsivity listed in the manual (i.e., not using the utensils provided was similar to not reading the instructions before starting the task). Functional ratings for each of the nine aspects of EF were not recorded during this initial training session.

During the second training session specific examples of cognitive activity limitations for directed and sustain attention, error detection, and error correction were described prior to the video replay of a second participant. This was done as the research assistant had asked for behavioural examples for these criteria. The research assistant also questioned the meaning of the statement in the scoring manual, under Initiation, "If a step is not completed, do not prompt the individual to initiate the uncompleted step after a lapse time of 1 minute has passed from the time the step should have been performed and do not score the uncompleted step". An explanation was given and the manual was subsequently modified for clarity. The sentence now reads, "If a step is skipped or omitted, do not prompt the individual to initiate the skipped or omitted step and do not score the uncompleted step". During the video replay, when the participant performed cognitive activity limitations, the videotape was paused and their performance was highlighted. After the completion of the videotape, the research assistant assigned functional ratings and capacity/incapacity scores. It was stressed that the Evidence for Capacity and Evidence for Incapacity scores were subjective and may not necessarily match the assigned functional rating.

During the third training session, the research assistant independently scored a third participant's performance in the presence of the trainer. After the video replay, functional ratings and capacity/incapacity scores were reviewed. Based on her rating of the participant's performance, the assistant was considered trained in the scoring method.

The research assistant then independently rated two participants on her own, followed by a review of her rating ability. It became apparent that the research assistant had misinterpreted one aspect of the scoring manual, that of how to score additional steps. Therefore, the trainer clarified that additional steps were only scored as Executive Memory errors if they detracted from the successful completion of the task. After scoring ten participants, the research assistant requested clarification about the Error Detection score. In particular, she queried whether or not every Planning and Organisation error should also be scored under Error Detection. She was referred to the manual which states that "examiners may believe that an undetected error has occurred, when in fact, the individual has simply organised the aspects of the task in an inappropriate manner". Therefore, she was instructed that Planning and Organisation errors are not automatically scored under Error Detection.

Procedure

The procedure used in the Main Assessment was described in the Developmental Trial (see Chapter Three, p. 64). Briefly, after obtaining informed consent, participants completed the functional task, the Activities of Daily Life Questionnaire, the EBIQ, the Demographic Questionnaire, the Sorting Test, and then the Tower Test. One formal memory measure was also administered to these participants, as part of data collection for a separate (though related) study that will be reported elsewhere. The one exception to the procedure described in Chapter Three was that the examiner scored participant performance solely based on videotaped footage for those participants in the Developmental Trial and solely on the live performance for new participants—a review of the videotaped footage was not necessary to clarify the scoring following any of the live performances. For participants with TBI, testing was conducted across two sessions lasting approximately one hour each, with these sessions separated across several days or weeks, whereas for participants in the normative group, testing was conducted in one session that lasted approximately two hours.

Planned Analysis

The inter-rater reliability of the method was assessed by way of ICC and was described in the Developmental Trial (see Chapter Three, p. 67). The dependent variables from the Activities of Daily Life Questionnaire and EBIQ were described in the Pre-Study Survey (see Chapter Two, p. 40). There were two sets of dependent variables from the functional task. The first, or objective rating, was the functional ratings of the nine aspects of EF, namely *Planning and Organisation*, *Executive Memory*, *Initiation*, *Cognitive Shifting*, *Impulsivity*, *Sustained and Directed Attention*, *Error Detection*, *Error Correction*, and *Time Management*, as described in the Developmental Trial (see Chapter Three, p. 69). The second set of dependent variables, or subjective ratings, were functional ratings of the nine aspects of EF that were derived from the subjective analysis (i.e., the *Evidence of Capacity* and *Evidence of Incapacity* scores).

The dependent variables from the Sorting Test were the scaled scores for confirmed correct sorts, free sorting description score, sort recognition description score, and combined description score. The dependent variables from the Tower Test were the scaled scores for total achievement score, mean first move time, move accuracy ratio, and total rule violations. These variables were calculated as per the instructions in Delis et al. (2001).

The directionality of the rating scales used in the Main Assessment were not consistent. For example, higher scores on the Sorting Test indicated better functioning, whereas lower scores on the functional task indicated better functioning. However, during data analysis all scales were modified such that higher scores indicated better performance (i.e., less disability or activity limitations) and lower scores indicated poorer performance (i.e., more disability or activity limitations). This was done so that any positive correlations would indicate a direct positive relationship, with any inverse correlations indicating an inverse relationship.

Pearson product-moment correlation coefficients were used to assess the relationship between variables of interest as this is the most widely used correlational coefficient. It is generally held that correlations between ± .01 and ± .29 indicate a weak relationship, correlations between ± .30 and ± .69 indicate a moderate relationship, and correlations above ± .70 indicate a strong relationship (Sheskin, 2004). However, research suggests that correlations between neuropsychological tests and everyday behaviours range from .24 to .50 (Burgess et al., 1998; Chaytor et al., 2006; Manchester et al., 2004). Therefore, the generally held view of the strength of relationships was modified in the Main Assessment. Specifically, correlations between ± .24 and ± .69 were considered moderate relationships.

Non-parametric tests were used in the analyses involving gender. Specifically, given that gender is a categorical variable, non-parametric tests were required during the analysis of this variable. Therefore, chi-squared tests of independence, Cramer's phi, and Spearman's rank correlation coefficients were utilised.

Finally, data obtained from participants in the TBI group were always analysed independently from the data obtained from participants in the normative group. In addition to this, analyses were also conducted after collapsing the two groups into one (i.e., during correlational analyses and a multiple regression analysis). There were two reasons for this. First, during correlational analyses participants act as their own control and, thus, each participant's performance was only compared with their own performance. Second, an adequate range of scores is required to detect relationships among variables. Therefore, combining the two groups had the effect of reducing the threat of range restriction.

Results

Inter-Rater Reliability

ICCs were used to determine the inter-rater reliability of the scoring method. According to Haneline's (2007) evaluation criteria, the ICCs for the *Planning and Organisation* (1.00), *Time Management* (1.00), *Executive Memory* (.78), and *Error Detection* (.77) scores were within the excellent range, whereas the ICCs for *Impulsivity* (.73) and *Sustained and Directed Attention* (.71) were within the fair to good range. The ICCs for the remaining three functional ratings, *Error Correction* (.35), *Cognitive Shifting* (.23), and *Initiation* (.00), were within the poor range.

Percentage agreements were calculated for three functional ratings in order to determine whether or not a reduced range of scores led to the ICC scores being in the poor range. Percentage agreement between raters for *Initiation*, *Cognitive Shifting*, and *Error Correction* were 95%, 86%, and 81%, respectively. These scores suggest that a reduction in range produced the low ICCs, rather than poor inter-rater agreement.

Comparison Between the Objective and Subjective Functional Ratings

In most cases, the objective and subjective functional ratings were strongly associated with each other; *Planning and Organisation*, $r = .88$, $p < .01$, *Executive Memory*, $r = .93$, $p < .01$, *Cognitive Shifting*, $r = .84$, $p < .01$, *Impulsivity*, $r = .81$, $p < .01$, *Error Detection*, $r = .95$, $p < .01$, *Error Correction*, $r = .84$, $p < .01$, and *Time Management*, $r = .95$, $p < .01$. *Sustained and Directed Attention* was the only functional rating in which a significant correlation was not found between the objective and subjective ratings; $r = .27$, $p = .08$. Given the significant association between the objective and subjective ratings, the objective ratings were used in all subsequent analyses.

Ecological Validity

Analysis with TBI group only. One of the aims of the Main Assessment was to examine the performance of people with TBI on the functional task compared to their performance on the formal measures of cognitive impairment. To this end, a correlational analysis was conducted between the functional task and scores on the Sorting Test and Tower Test for participants with TBI (see Table 9). Each participant received an *Initiation* score of Full Function, and, therefore, this variable was removed from the analysis. Significant correlations were found between scores on the functional task and the Sorting Test. In particular, *Planning and Organisation* was significantly correlated with Confirmed Correct Sorts, $r = .47$, $p < .05$, Free Sorting Description Score, $r = .48$, $p < .05$, and Combined Description Score, $r = .46$, $p < .05$, *Impulsivity* was significantly correlated with Confirmed Correct Sorts, $r = .47$, $p < .05$, the Free Sorting Description Score, $r = .46$, $p < .05$, Sort Recognition Description Score, $r = .43$, $p < .05$, and Combined Description Score, $r = .51$, $p < .05$, and *Error Correction* was significantly correlated with Confirmed Correct Sorts, $r = .41$, $p < .05$, the Sort Recognition Description Score, $r = .47$, $p < .05$, and Combined Description Score, $r = .46$, $p < .05$. In addition to these significant correlations, *Executive Memory*, *Cognitive Shifting*, *Error Detection*, and *Time Management* yielded moderate, non-significant, correlations with the Sorting Test. The only functional rating that did not yield either significant or moderate, non-significant, correlations with the Sorting Test was *Sustained and Directed Attention* (see Table 9).

The only significant correlation between the functional task and the Tower Test was between *Impulsivity* and the Total Achievement Score, $r = .51$, $p < .01$. However, moderate, non-significant, correlations were found between *Planning and Organisation*, *Executive Memory*, *Error Detection*, *Error Correction*, and *Time Management* and the Tower Test. *Cognitive Shifting* and *Sustained and Directed Attention* yielded neither significant nor moderate, non-significant, correlations with the Tower Test (see Table 9).

A number of significant inverse correlations were found between the functional task and scores on the EBIQ. For example, the functional rating *Impulsivity* yielded a significant, inverse, correlation with self-reported impulsivity as measured by the EBIQ, $r = -.49$, $p < .05$ (see Table 10). The only significant correlation between the functional task and family member ratings on the EBIQ was between *Executive Memory* and the Depression sub-scale, $r = .63$, $p < .05$. However, a number of non-significant correlations were found between the two measures (see Table 11). There were no significant differences between participant self-report or family members rating on the EBIQ (see Table 12).

For participants with TBI, significant correlations were found between self-reported ability to make a simple meal and the functional ratings *Planning and*

Organisation, $r = .39, p < .05, df = 24$, *Cognitive Shifting*, $r = .47, p < .05, df = 24$, *Error Correction*, $r = .83, p < .01, df = 24$, and *Time Management*, $r = .47, p < .05, df = 24$. In addition, a significant correlation was found between self-reported ability to make a complex meal and *Error Correction*, $r = .51, p < .01, df = 24$, for participants with TBI. Each participant, and their family member, reported that they could make a cup of coffee or tea independently. There were no significant differences between participant and family-member ratings on the Activities of Daily Life Questionnaire.

Analysis with all participants. Given the associations between the functional task and the formal neuropsychological measures in the TBI group, an analysis utilising all participants in the Main Assessment was undertaken. A number of significant correlations were found (see Table 13). For example, *Planning and Organisation* was significantly associated with Confirmed Correct Sorts, $r = .41, p < .01$, and the Combined Description Score, $r = .45, p < .01$, from the Sorting Test, and the Total Achievement Score, $r = .36, p < .05$, from the Tower Test. *Executive Memory* was significantly associated with Confirmed Correct Sorts, $r = .33, p <.05$, and the Combined Description Score, $r = .35, p < .05$, from the Sorting Test. *Impulsivity* was significantly associated with Confirmed Correct Sorts, $r = .49, p < .01$, and the Combined Description Score, $r = .59, p < .01$, from the Sorting Test, and the Total Achievement Score, $r = .47, p < .01$, from the Tower Test. *Error Detection* was significantly correlated with the Combined Description Score from the Sorting Test, $r = .39, p < .05$, and the Total Achievement Score, $r = .39, p < .05$, from the Tower Test. *Error Correction* was significantly related to Confirmed Correct Sorts, $r = .41, p < .01$, and the Combined Description Score, $r = .45, p < .01$, from the Sorting Test, whereas *Time Management* was significantly related to the Free Sorting Description Score,

$r = .33, p < .05$, from the Sorting Test and to Total Rule Violations, $r = .45, p < .01$, from the Tower Test. In addition to these significant correlations, a number of moderate, non-significant associations were also found (see Table 13).

Several sub-scales of the EBIQ were associated with the functional task. However, on measures that one might expect to find a significant association, such as between the EBIQ sub-scale Cognitive and the functional ratings, there were no such significant correlations. In many instances, the analysis yielded non-significant, inverse, correlations (see Table 14).

Comparison Between Groups

Another aim of the Main Assessment was to compare the performance of the group with TBI and the normative group on the functional task. The normative group performed significantly better than the TBI group on *Planning and Organisation*, $t = 2.27, p < .01, d = .72$, *Executive Memory*, $t = 3.91, p < .01, d = 1.23$, *Cognitive Shifting*, $t = 1.86, p < .05, d = .59$, *Impulsivity*, $t = 1.86, p < .01, d = 1.19$, *Error Detection*, $t = 3.45, p < .05, d = 1.09$, and *Time Management*, $t = 1.78, p < .05, d = .56$. Group differences on *Sustained and Directed Attention*, $t = 1.45, p = .08, d = .46$, approached significance, whereas group differences on *Error Correction* were neither significant, nor approached significance (see Table 15).

The groups also differed significantly on the EBIQ, with participants with TBI reporting experiencing significantly more difficulties on all of the subscales than participants in the normative group; Somatic, $t = 3.02, p < .01, d = .96$, Cognitive, $t = 4.68, p < .01, d = 1.49$, Motivation, $t = 2.76, p < .01, d = .88$, Impulsivity, $t = 3.19$, $p < .01, d = 1.01$, Depression, $t = 3.73, p < .01, d = 1.18$, Isolation, $t = 3.26, p < .01$, $d = 1.03$, Physical, $t = 3.17, p < .01, d = 1.01$, Communication, $t = 3.16, p < .01$,

$d = 1.00$, and Core, $t = 5.04$, $p < .01$, $d = 1.60$ (see Table 16).

Demographic Data

Demographic characteristics between the two groups were analysed to determine the degree to which they played a role in the group differences found on the functional task. The groups did not differ significantly with regard to gender, $x^2 = 1.17$, $p > .05$, $\Phi = .17$, education, $t = .74$, $p > .05$, $d = .24$, baking experience, $t = -1.26$, $p > .05$, $d = .41$, confidence in baking, $t = -1.22$, $p > .05$, $d = .40$, or their satisfaction with the outcome of the task, $t = -1.52$, $p > .05$, $d = .48$. When asked to rate how well they thought they performed on the functional task, participants in the normative group rated their performance significantly higher than did participants in the TBI group, $t = -2.20$, $p < .05$, $d = .69$.

For the TBI group, gender was significantly related to *Executive Memory*, $\rho = .47$, $p < .05$, *Impulsivity*, $\rho = .52$, $p < .01$, and *Error Detection*, $\rho = .52$, $p < .01$. In addition, females performed significantly better than males did on *Executive Memory,* $t = 2.54$, $p < .05$, $d = 1.03$, *Impulsivity*, $t = 3.08$, $p < .01$, $d = 1.25$, and *Error Detection*, $t = 2.96$, $p < .01$, $d = 1.21$. The only other other demographic variables to be related to the functional ratings for the TBI group was the significant correlation between education and *Cognitive Shifting*, $r = .41$, $p < .04$, and between age and *Error Correction*, $r = -.41$, $p < .05$.

For the normative group, gender was only significantly related to *Executive Memory,* $\rho = .52$, $p < .05$, with women performing better than men on this functional rating, $t = 4.27$, $p < .05$, $d = 1.14$. In addition, baking experience was significantly associated with *Sustained and Directed Attention*, $r = .51$, $p < .05$, confidence in baking was significantly related to *Sustained and Directed Attention*, $r = .50$, $p < .05$, and self-reported performance on the task was significantly correlated with *Planning and Organisation*, $r = .53$, $p < .03$.

A multiple regression analysis revealed that participant group and gender made up the majority of the accounted-for variance for *Executive Memory*, $R^2 = .44$, *Impulsivity*, $R^2 = .37$, and *Error Detection*, $R^2 = .39$, whereas gender alone accounted for the majority of the accounted-for variance for *Sustained and Directed Attention*, $R^2 = .22$. Baking experience and confidence in baking did account for much of the variance in performance on these variables (see Table 17). Group, gender, baking experience, and confidence in baking were not significant factors in *Planning and Organisation*, *Initiation*, *Cognitive Shifting*, *Error Correction*, and *Time Management*.

Table 9

Pearson's Correlations Between the Functional Task and the Sorting and Tower Tests for Participants with TBI (n = 27).

Variables	1	2	3	4	5	6	7	8
Functional Task								
1. Planning and Organisation	—							
2. Executive Memory	.38*	—						
3. Cognitive Shifting	.34	.54**	—					
4. Impulsivity	.50**	.40*	.30	—				
5. Sustained and Directed Attention	.04	.17	.48*	-.05	—			
6. Error Detection	.44*	.79**	.25	.45*	-.11	—		
7. Error Correction	.38*	.29	.40*	.34	.12	.16	—	
8. Time Management	.28	.28	.56*	.16	.17	.02	.48*	—
Sorting Test								
Confirmed Correct Sorts[a]	.47*	.31	.25	.47*	.13	.23	.41*	.36
Free Sorting Description Score[a]	.48*	.29	.24	.46*	.04	.24	.37	.37
Sort Recognition Description Score[a]	.30	.15	.23	.43*	.00	.26	.47*	.26
Combined Description Score[a]	.46*	.30	.29	.51**	-.06	.30	.46*	.35
Tower Test								
Total Achievement Score[a]	.33	.27	.06	.51**	-.11	.28	.22	.01
Mean First-Move Time[a]	-.02	-.07	.01	-.26	-.02	-.28	.00	.20
Move Accuracy Ratio[a]	.22	.03	.15	-.20	.22	.07	.11	.23
Total Rule Violations[a]	.37	.13	.00	.19	-.03	-.02	.35	.34

Note. [a] n = 25, * $p < .05$, ** $p < .01$

Table 10

Correlations Between the Functional Task and Self-Reported EBIQ Scores for Participants with TBI (n = 26)

Variables	A	B	C	D	E	F	G	H
Somatic	-.40*	-.58**	-.42*	-.50**	-.05	-.54**	-.16	-.47*
Cognitive	-.42*	-.69**	-.55**	-.36	-.12	-.54**	-.11	-.46*
Motivation	-.23	-.22	-.27	-.28	.03	-.04	-.14	-.12
Impulsivity	-.51**	-.46*	-.44*	-.48**	-.06	-.35	-.38*	-.45*
Depression	-.02	-.13	-.08	-.39*	.18	-.18	.05	-.21
Isolation	-.23	-.36	-.33	-.32	.04	-.35	-.24	-.39*
Physical	-.16	-.28	-.12	-.14	.00	-.22	.10	-.16
Communication	-.20	-.34	-.33	-.34	-.07	-.23	.12	-.03
Core	-.31	-.44*	-.39*	-.45*	.02	-.37	-.09	-.39*

Note. A = planning and organisation; B = executive memory; C = cognitive shifting; D = impulsivity; E = sustained and directed attention; F = error detection; G = error correction; H = time management. Initiation was not included in the analysis as the variable was a constant.

$^{*}p < .05$, $^{**}p < .01$

Table 11

Correlations Between the Functional Task and Family Member Reported EBIQ Scores for Participants with TBI (n = 15)

Variables	A	B	C	D	E	F	G
Somatic	-.36	-.18	-.14	-.08	.02	-.17	-.45
Cognitive	-.36	-.30	-.20	.01	.08	-.16	-.38
Motivation	-.07	.40	.12	.23	.00	.42	-.16
Impulsivity	-.14	.27	.11	.22	-.12	.35	-.32
Depression	.32	.63*	.34	-.06	.36	.44	-.20
Isolation	-.29	.32	.25	-.02	.30	.13	-.36
Physical	-.25	-.04	-.14	.10	-.04	.02	-.27
Communication	-.36	-.24	-.25	-.11	-.08	-.21	-.21
Core	-.14	.30	.12	.10	.16	.27	-.41

Note. A = planning and organisation; B = executive memory; C = cognitive shifting; D = impulsivity; E = sustained and directed attention; F = error detection; G = time management. Initiation and Error Correction were not included in the analysis as one of the variables was constant.

$^* p < .05$

Table 12

Comparisons Between Self-Rating and Family Member-Rating for Participants with TBI on the EBIQ Subscales

	TBI Group (*n* = 27)		Family Member Group (*n* = 14)				
Variables	M	SD	M	SD	t	p	d
Somatic	1.76	.57	1.81	.41	-.28	.39	.09
Cognitive	1.74	.50	1.69	.47	.34	.37	.11
Motivation	1.50	.54	1.51	.51	-.06	.48	.02
Impulsivity	1.60	.50	1.74	.60	-.75	.23	.25
Depression	1.50	.49	1.63	.43	-.80	.21	.26
Isolation	1.69	.49	1.88	.47	-1.13	.13	.37
Physical	1.36	.45	1.49	.45	-.88	.19	.29
Communication	1.62	.56	1.59	.43	.18	.43	.06
Core	1.61	.43	1.70	.38	-.65	.26	.21

Table 13

Correlation Matrix Between the Functional Task and the Sorting and Tower Tests for All Participants (n = 43)

Variables	1	2	3	4	5	6	7	8
Functional Task								
1. Planning and Organisation	—							
2. Executive Memory	.39**	—						
3. Cognitive Shifting	.36*	.54**	—					
4. Impulsivity	.53**	.53**	.39*	—				
5. Sustained and Directed Attention	.14	.33*	.45*	.08	—			
6. Error Detection	.47**	.83**	.34*	.56**	.12	—		
7. Error Correction	.37*	.32*	.43**	.38*	.14	.22	—	
8. Time Management	.36*	.38*	.53**	.26	.33*	.17	.46**	—
Sorting Test								
Confirmed Correct Sorts[a]	.41**	.33*	.28	.49**	.11	.28	.41**	.29
Free Sorting Description Score[a]	.44**	.36*	.29	.51**	.11	.34*	.37*	.33*
Sort Recognition Description Score[a]	.35*	.25	.31*	.54**	.00	.36*	.44**	.23
Combined Description Score[a]	.45**	.35*	.35*	.59**	.08	.39*	.45**	.30
Tower Test								
Total Achievement Score[a]	.36*	.30	.11	.47**	.09	.39*	.21	.07
Mean First-Move Time[a]	.03	.01	.03	.12	.01	-.15	.01	.12
Move Accuracy Ratio[a]	.13	.07	.11	.24	.14	.00	.08	.16
Total Rule Violations[a]	.43**	.20	.05	.23	.12	.12	.31*	.45**

Note. [a] n = 41, *p < .05, ** p < .01.

Table 14

Correlations Between the Functional Task and Self-Reported EBIQ Scores for All Participants (n = 41)

Variables	A	B	C	D	E	F	G	H
Somatic	-.21	-.01	-.07	-.14	-.04	-.13	.06	-.08
Cognitive	-.15	.04	-.05	.12	-.03	-.02	.15	-.05
Motivation	-.04	.27	.09	.06	.11	.32*	.05	.23
Impulsivity	-.27	.08	-.14	-.13	-.07	.09	-.17	-.07
Depression	-.16	.43*	.32*	.01	.18	.24	.28	.15
Isolation	.00	.21	.11	.07	.13	.10	-.02	.09
Physical	.01	.22	.20	.18	-.03	.17	.33*	.20
Communication	-.04	.19	.03	.03	-.08	.13	.32*	.25
Core	-.06	.25	.07	.03	.06	.15	.19	.08

Note. A = planning and organisation; B = executive memory; C = cognitive shifting; D = impulsivity; E = sustained and directed attention; F = error detection; G = error correction; H = time management. Initiation was not included in the analysis as the variable was a constant.

$^*p < .05$

Table 15

Comparison Between the Functional Ratings for the TBI and Normative Groups

	TBI Group (*n* = 27)		Normative Group (*n* = 16)				
Variables	M	SD	M	SD	t	p	d
Planning and Organisation	1.74	.66	1.31	.48	2.27	.01	.72
Executive Memory	2.15	.77	1.31	.48	3.91	.00	1.23
Initiation	1.00	.00	1.00	.00			
Cognitive Shifting	1.19	.40	1.00	.00	1.86	.04	.59
Impulsivity	1.48	.51	1.00	.00	3.76	.00	1.19
Sustained and Directed Attention	1.33	.48	1.13	.34	1.52	.08	.46
Error Detection	2.67	1.41	1.38	.62	3.45	.00	1.09
Error Correction	1.22	.70	1.00	.00	1.27	.11	.40
Time Management	1.74	1.10	1.19	.75	1.78	.04	.56

Table 16

Comparison Between EBIQ Scores for the TBI and Normative Groups

	TBI Group (*n* = 26)		Normative Group (*n* = 16)				
Variables	M	SD	M	SD	t	p	d
Somatic	1.83	.46	1.45	.26	3.02	.00	.96
Cognitive	1.81	.36	1.32	.26	4.68	.00	1.49
Motivation	1.56	.46	1.21	.26	2.76	.00	.88
Impulsivity	1.67	.39	1.33	.20	3.19	.00	1.01
Depression	1.56	.40	1.17	.17	3.73	.00	1.18
Isolation	1.76	.36	1.41	.30	3.26	.00	1.03
Physical	1.41	.36	1.09	.21	3.17	.00	1.01
Communication	1.68	.47	1.25	.37	3.16	.00	1.00
Core	1.67	.29	1.26	.19	5.04	.00	1.60

Table 17

Multiple Regression Analysis of Variables in Relation to the Functional Task (n = 43)

Variable	*B*	*SE B*	β	*t*	*p*
			Executive Memory		
Intercept	3.36	.49		6.88	.00
Group	-.69	.19	-.47	-3.57	.00
Sex	-.70	.23	-.48	-2.97	.00
Experience	.03	.12	.04	.20	.84
Confidence	.09	.09	.17	1.02	.32
			Impulsivity		
Intercept	2.25	.32		6.96	.00
Group	-.40	.13	-.44	-3.15	.00
Sex	-.33	.16	-.37	-2.15	.04
Experience	-.03	.08	-.09	-.42	.68
Confidence	.04	.06	.20	1.11	.28
			Sustained and Direct Attention		
Intercept	1.79	.36		4.96	.00
Group	-.17	.14	-.18	-1.18	.25
Sex	-.39	.17	-.44	-2.28	.03
Experience	.06	.09	.15	.65	.52
Confidence	-.04	.07	-.11	-.53	.60
			Error Detection		
Intercept	4.76	.91		5.25	.00
Group	-1.08	.36	-.41	-3.03	.01
Sex	-1.09	.43	-.42	-2.51	.02
Experience	-.11	.23	-.10	-.49	.63
Confidence	.23	.17	.24	1.37	.18

Note. Experience = baking experience; Confidence = confidence in baking.

Discussion

One of the purposes of the Main Assessment was to examine the performance of people with TBI on the functional task compared to their performance on formal neuropsychological measures of EF. Both significant and moderate, non-significant, correlations between the functional "real-life" measure and scores on the Sorting Test and Tower Test were found. This is consistent with past studies that have found moderate correlations, ranging from r = .24 to .50, between formal neuropsychological measures and ecologically valid measures of EF (Burgess et al., 1998; Chaytor et al., 2006; Manchester et al., 2004).

As discussed in Chapter One, EF represent a group of high-level cognitive processes. Herein lies one of the difficulties in assessing EF. In the brief review earlier, 11 different aspect of EF were identified from the literature (see Table 1). Nine of these aspects were assessed in the functional task and compared with performance on two formal neuropsychological measures—the Sorting Test and the Tower Test. Although the results found that most of the functional ratings yielded moderate correlations with the formal measures, Sustained and Directed Attention did not. One of the reasons for this finding may be that neither the Sorting nor the Tower Tests specifically assess sustained or directed attention (see Delis et al., 2001). Each participant in the Main Assessment also received an Initiation of *Full Function*. Therefore, although the formal neuropsychological measures were designed to identify people with initiation difficulties, the functional task may not have been sensitive enough to identify problems in this area.

The functional task was found to display good to excellent inter-rater reliability. Although some of the functional ratings yielded a restricted range of scores, inter-rater agreement among these functional ratings was high. These results replicate those found in the Developmental Trial and support the reliability of the functional task. In addition, it required a total of four hours to train the second rater to produce accurate scores. This is quite a brief training period when compared to other behavioural observations measures. For example, it requires a five-day training workshop in order to be qualified to score the AMPS (Bernspång, 1995; K. Sullivan, personal communication, September 24, 2009), 15 hours of training was conducted by raters before scoring the Pre-Linguistic Autism Diagnostic Observation Schedule (DiLavore, Lord, & Rutter, 1995), and six hours of training are required to become trained in using the FIM (Quinsey, Findlay, & Willmott, 2005).

Performance on the functional task was also compared to scores on the EBIQ. Although no formal hypothesis was made with regard to the strength of this relationship, it was expected that performance on the functional task would be associated with the EBIQ subscales that tap cognitive aspects of functioning (i.e., Cognitive, Motivation, Impulsivity, and Core). For participants with TBI, their self-report scores on the EBIQ were largely inversely correlated with performance on the functional task. For example, the Cognitive subscale of the EBIQ yielded significant, inverse, correlations with the functional ratings Planning and Organisation, Executive Memory, Cognitive Shifting, Error Detection, and Time Management and the Impulsivity subscale of the EBIQ yielded a significant, inverse, correlation with the functional rating Impulsivity.

Similar results were also found when performance on the functional task was compared with family member-ratings on the EBIQ. However, in a number of instances, the associations were in the expected direction (e.g., the correlation between the functional rating Executive Memory and the Core subscale was r = .30). Other studies have similarly found that performance on functional tasks is not significantly associated with scores on questionnaires designed to assess cognitive and emotional functioning (see Chevignard et al., 2000, 2009). There may be several reasons for these findings. First, there are inherent problems with using questionnaires with people who have sustained a TBI. For example, self-awareness difficulties (Hart et al., 2005) and a number of proxy factors, such as intra-familial distress and stereotypes (Gioia et al., 2008; Manchester et al., 2004), affect the accuracy of information gathered. Indeed, Chevignard et al. (2000) found that when information was obtained from occupational therapists, significant correlations emerged between performance on their cooking task and scores on the Dysexecutive Questionnaire. Second, a number of participants in the current study were residing at an inpatient rehabilitation centre when they completed the testing. Therefore, participants and their family members may not have been able to accurately answer some of the questions on the EBIQ, such as ‘Problems with household chores’ or ‘Lack of interest in hobbies at home’. Third, the validity of the EBIQ was initially established as it could be used to discriminate people who had a TBI from people who had not sustained a TBI (Teasdale et al., 1997)—a capability that is supported by the current study. However, concurrent validity was not assessed (Deb et al., 2007). Holm, Schönberger, Poulsen, and Caetano (2009) found that when they compared scores on the EBIQ to scores on the FIM, another commonly used outcome measure in brain injury rehabilitation, self-report scores on the Cognitive subscale and family member-report scores on the Physical subscale were the only EBIQ subscales that yielded significant correlations with the total FIM score. Therefore, further research into the psychometric properties of the EBIQ is needed to establish the validity of this measure.

Despite these findings, the results from the EBIQ are consistent with other research. Specifically, the Pre-Study Survey, the Main Assessment, and Svendsen et al.

(2004) found that self-report ratings from people with TBI did not differ significant from ratings by family members. In addition, results from the Main Assessment indicate that participants with TBI reported significantly more difficulties than did participants in the normative group, a finding that is supported by Teasdale et al. (1997).

One of the difficulties researchers have in developing ecologically valid measures of EF is that there is no ‘gold standard’ measure with which to compare any newly developed measure. Although the EBIQ was used in the current study as the ecologically valid comparator, the Dysexecutive Questionnaire is more frequently used in this capacity (see Burgess et al., 1998; Chaytor et al., 2006; Chevignard et al., 2000, 2008, 2009; Crawford & Henry, 2005; Gioia et al., 2008; Hart et al., 2005; Knight et al., 2002). One criticism of the Main Assessment may be that the Dysexecutive Questionnaire should have been used, rather than the EBIQ. However, Wood and Liossi (2006) found that the Dysexecutive Questionnaire failed to yield significant correlations with several subtests from the BADS. Given that the BADS was specifically designed to be ecologically valid (see Chamberlain, 2003), Wood and Liossi state that the Dysexecutive Questionnaire should also not be used as a gold standard measure of ecological validity until further research is undertake to clarify the psychometric properties of the measure.

A second questionnaire, used to gather information about each participant’s level of independence in completing cooking tasks, was also administered in the Main Assessment. As previously discussed, the ecological validity of a particular measure is dependent on the operational definition of everyday behaviour (see Chaytor et al., 2007). Given that the functional task was a cooking task, one would expect that cooking ability would be associated with the functional task. Indeed, significant associations were found between performance on the functional task and self-reported ability to make simple and complex meals. Thus, these results provide initial support for the validity of the functional task as an ecologically valid measure of cognitive functioning.

Another aim of the Main Assessment was to compare the performance of people with TBI to people without brain injuries on the functional task. Participants with TBI displayed significantly more cognitive activity limitations than did participants in the normative group on six of the nine functional ratings, with another functional rating approaching significance. Therefore, the functional task can be used to discriminate between those who have sustained a TBI and those who have not.

Cooking Tasks as Measures of EF

A cooking task was chosen as the functional task based on information gathered from the Pre-Study Survey. Although participants in the Pre-Study Survey did not specifically identify that they make chocolate brownies, this specific task was chosen because it mirrored previous research (see Baguena et al., 2006; Chevignard et al., 2000, 2009), utilised a microwave oven, and because it was thought that participants would enjoy making a chocolate treat. However, one of the difficulties in developing ecologically valid measures of cognitive functioning is that the validity of any measure is dependent on the operational definition of everyday functioning (Chaytor et al., 2007). Thus, although the current results provide support for the utility of the measure, it may only be useful in assessing executive functions as they pertain to cooking, or more specifically, to chocolate brownie making.

One of the participants in the study stated that they had never made brownies before, would never made chocolate brownies in the future, and if they ever wanted them, they would simply buy them at a grocery store. This participant also stated that they cooked simple meals that required few ingredients. The functional task for this participant, and other people like him, is unlike any everyday behaviour that they perform. The purpose of ecologically valid measures of cognitive functioning is to provide a prediction about how well someone will complete a particular everyday task (Gioia & Brekke, 2009). For the participant just described, the functional task may underestimate his ability to complete everyday cooking tasks because of the simplicity of the meals he makes for himself. Furthermore, it would likely underestimate his ability to produce an edible brownie at a social gathering, as he would have simply purchased the brownie at a grocery store—provided his prospective memory is such that he remembers to purchase the brownie prior to the social engagement. Thus, the complexity and specifics of the functional task may reduce its ecological validity for people like this participant (see Chaytor et al., 2007).

Females in the TBI group performed significantly better than did males with TBI on the functional ratings Executive Memory, Impulsivity, and Error Detection, whereas females in the normative group only performed significantly better than did males in the normative group on the functional rating Executive Memory. Research indicates that the increased levels of oestrogen and progesterone found in females are protective factors following TBI (Groswasser, Cohen, & Keren, 1998; Ottochian et al., 2009; Stein, 2001). These biological differences may be one reason why females performed better than males in the TBI group (i.e., the females in the group experienced less executive dysfunction than did males due to a physiological advantage). Males in the normative group were expected to perform similarly to females as they had not experienced a TBI. Indeed, this occurred for eight of the nine functional ratings. Other factors may have contributed to the gender differences found within the groups. For instance, cooking, which has traditionally been a female task (Fürst, 1997), may have been easier for females than for males because of the reduced novelty of cooking tasks. In addition, female participants' increased experience in cooking may have caused the functional task to be more habitual in nature. This would introduce the use of other cognitive processes, other than EF, during the completion of the task. The results of the current study also found that gender accounted for a significant portion of the variance in Executive Memory, Impulsivity, Sustained and Directed Attention, and Error Detection scores. Thus, any measure that uses cooking to assess cognitive functioning with people who have sustained a TBI will need to take into account the inherent confound that gender presents. One way to control for gender differences during functional tasks may be to develop normative information, or cut-off criteria, specifically for males and females, such as those developed for traditional neuropsychological measures (see Strauss, Sherman, & Spreen, 2006).

Several studies have attempted to create ecologically valid measures using cooking tasks (see Baguena et al., 2006; Baum & Edwards, 1993; Chevignard et al., 2000, 2008, 2009; Neistadt, 1994; Schwartz et al., 2002). However, Neistadt and Baum and Edwards did not compare their measures against formal measures of EF and Schwartz et al. only compared their cooking task to formal measures of attention, although sustained attention was one of the aspects of attention assessed by the formal neuropsychological measures. The studies by Chevignard et al. did not find significant correlations between their cooking task and formal measures of EF, whereas the study by Baguena et al., like the current study, did find significant correlations between their functional measure and formal measures of EF.

One of the criticisms of the studies that compared their cooking task to formal measures of EF is that they only reported statistically significant associations between the measures. Research suggests that moderate correlations exist between formal neuropsychological measures and ecologically valid measures of cognitive functioning.

Given this finding, one would not necessarily expect to find significant correlations between cooking tasks and formal measures of EF, but would expect to find moderate ones. Only relying on and reporting statistically significant associations may result in an underestimation of the ecological validity of the functional measure, depending on the sample size of the study. Therefore, any future research that attempts to establish the ecological validity of a measure should report non-significant, moderate, correlations between the everyday measure and formal neuropsychological measures, focusing on effect sizes rather than statistical significance which is irrevocably associated with sample size rather than importance. Depending of the sample size, even non-significant associations may be viewed as preliminary evidence for the ecological validity of a newly developed measure.

Several behavioural observation measures for assessing everyday EF have been developed over the past two decades. A number of these, which involve route-finding and errand completion, rely heavily on perceptual and spatial abilities (see Boyd & Sautter, 1993). Thus, everyday tasks which do not involve navigation may prove to be more valid measures of EF. Cooking tasks, which do not require navigation, are well placed in this respect. In the Main Assessment, a number of participants were able to complete the task while seated in a chair. However, perceptual and spatial abilities may be an inherent confound in any functional task (e.g., participants in the current study were required to scan their environment in order to select the correct ingredient).

Is the Functional Task a Suitable Outcome Measure?

Turner-Stokes (2002) reported a series of criteria for establishing the appropriateness of outcome measures in brain injury rehabilitation. She states that first, they need to possess adequate reliability and validity. In the current study, the functional task yielded significant and non-significant, moderate, correlations with formal measures of EF. The functional task was also able to discriminate between participants with TBI and participants in the normative group. Although these are results from an initial trial, they provide preliminary support for the validity of the functional task, and more generally in the approach taken in developing and scoring the particular task used. Furthermore, the inter-rater reliability of the measures was established in both the Developmental Trial and the Main Assessment.

Second, outcome measures need to be able to assess change as a result of completing a rehabilitation programme. When participants embarked on the functional task, they were instructed that they could utilise any type of strategy to help them complete the task. In this way, if people are taught compensatory strategies in a rehabilitation programme, they are allowed to utilise the strategies in helping them complete the task. Furthermore, the use of these strategies does not affect the scores people receive on the task. For example, one participant with TBI sustained damage to his hand as a result of his accident. During the functional task, this participant requested that the examiner break the eggs into the mixture. He requested that this step occur at the correct stage in the task and, thus, was not scored as having performed an error. This example highlights how the use of rehabilitation strategies can be used during the functional task. Although the current study did not assess performance before and after a cognitive rehabilitation programme, the functional task should be sensitive to changes in functioning that arise following cognitive rehabilitation.

Third, Turner-Stokes (2002) explains that outcome measures need to have a single set of scoring criteria, described in a manual. Two manuals were created in the Developmental Trial and Main Assessment. The Administration Manual describes how to set-up the functional task (e.g., place all ingredients and utensils on a table or bench) and how to respond to questions posed by participants (e.g., if participants ask for the examiner to confirm that what they are doing is correct, the examiner should answer their questions). The Scoring Manual provides a single set of criteria necessary to score participant performance.

Fourth, outcomes measures must have the potential to be readily incorporated into clinical practice. This has implications for how and where the measure is administered and the time needed to complete the measure (Turner-Stokes, 2002). One of the reasons for developing a task that utilised a microwave oven was so that the task could be easily transportable. Although the majority of participants completed the task in a kitchen setting, some participants completed the task in rooms with no kitchen facilities. In these latter instances, a microwave and kettle were brought into these rooms and placed on a table next to the ingredients. The ability to conduct testing in virtually any environment adds to the practicality of the measure. The transportability of the current functional task contrasts other functional measures, such as the ones described in Baum and Edwards (1993) and Chevignard et al. (2000, 2008, 2009), in which the tasks needed to be performed in kitchens as the actual cooking of the food required a stove or oven.

In the current study, participants were asked to stop the functional task if they had not completed the task within 40 minutes. In many instances, participants with TBI completed the task in less than 30 minutes. Although this is longer than the 10 to 15 minutes it generally took to complete the Tower and Sorting Tests, it is shorter than the near 80 minutes to generally took for participants with TBI to complete the cooking task in Chevignard et al.'s (2000) study. Thus, the relative brevity of the current functional task has the potential to be readily incorporated as part of an assessment that also involves standard psychometric assessment measures.

Fifth, Turner-Stokes (2002) states that measures that are already established outcome measures are more likely to be utilised than newly developed measures as clinicians may be hesitant to introduce new measures into their practice. The functional task in the current study does not meet this criterion as it has been recently created and only trialled for the purposes of this research. However, all new measures fail to meet this criterion and if a measure is developed that is found to be useful in assessing rehabilitation outcomes, clinicians may introduce the measure into their general practice.

The sixth criterion, an electronic system to score and collate data (Turner-Stokes, 2002), has only partially been completed. Although a spreadsheet was created using Microsoft Excel to convert the number of instances of cognitive activity limitations performed to functional ratings, no formal system has yet been created to collate the data. Finally, Turner-Stokes stated that any training required to properly administer and score the measure needs to be inexpensive. As described above, four hours of training were required for a second rater to be considered trained in the scoring procedure. This is less time than it takes to train raters for other behavioural observation methods (see Bernspång, 1995; DiLavore et al., 1995) and may be seen as a more economically viable alternative than a measure that requires numerous hours of training. Although the functional task has only recently been developed and trialled, many of the criteria described by Turner-Stokes have been met. Therefore, this functional task may be seen as a suitable outcome measure in brain injury rehabilitation.

Verisimilitude vs Veridicality

Ecologically valid measures of cognitive functioning are identified using one of two methods—verisimilitude and veridicality (Chaytor & Schmitter-Edgecombe, 2003).

Verisimilitude refers to the extent to which a particular measure appears to reflect the cognitive processes necessary to complete an everyday activity (Wood & Liossi, 2006). For example, a task using a driving simulator has high verisimilitude to driving ability, whereas the verisimilitude between driving ability and card sorting would be low. Veridicality, on the other hand, refers to the extent to which an existing measure is associated with everyday functioning (Wood & Liossi). This latter method is utilised when formal measures of cognitive functioning are compared with ecologically valid measures.

Cooking tasks, such as the one developed in the current series of studies, are examples of measures designed using the verisimilitude approach—the cognitive processes necessary to complete cooking tasks in the laboratory are presumed to be similar to the cognitive processes needed to complete cooking tasks when outside the laboratory. Chaytor and Schmitter-Edgecombe (2003) explain that one of the limitations of measures designed with verisimilitude is that they may not be scientifically rigorous. This is because these measures only need to appear to assess similar everyday cognitive processes, rather than actually assessing similar cognitive processes. However, this does not imply that all verisimilitude measures lack scientific rigour. In the Main Assessment, the functional task was shown to possess adequate reliability and validity.

Several authors, including Chaytor and Schmitter-Edgecombe (2003), Piotrowski and Keller (1989), and Turner-Stokes (2002), have reported that neuropsychologists are hesitant to incorporate newly developed measures into their clinical practice. Therefore, clinicians may not utilise these newly developed measures. This may be one reason for the numerous studies which have assess the ecological validity of formal neuropsychological measures (see Burgess et al., 1998; Chaytor et al., 2006; Hoskin, Jackson, & Crowe, 2005; Manchester et al., 2004; Mitchell & Miller, 2008; Odhuba et al., 2005; Ready et al., 2001; Semkovska, Bedard, Godbout, Limoge, & Stip, 2004; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2008; Wood & Liossi, 2006). Despite this extensive research, the ecological validity of formal measures tends to remain in the low to moderate range.

Predicting someone's ability to perform everyday tasks should not be solely based on their performance on neuropsychological tests. Indeed, in current neuropsychological practice, clinicians tend to use a variety of methods to determine their client's ability to perform everyday tasks. These methods include formal neuropsychological measures (veridicality), questionnaires, behavioural observations, clinical interviews, and medical records (Sbordone & Guilmette, 1999). However, in their study of commonly used neuropsychological measures, Rabin et al. (2007) found that less than 10% of respondents reported that they frequently used ecologically valid measures (verisimilitude) of memory, attention, or EF. Thus, one important avenue for information gathering appears to be being neglected.

Wilson (2000) stated that the purpose of cognitive rehabilitation is to provide people with the skills, or alter the environment sufficiently, to maximise independence in daily life. Thus, obtaining accurate information about someone's current everyday functioning is critical if cognitive rehabilitation is to be implemented effectively. Furthermore, this information is also important in assessing the effectiveness of cognitive rehabilitation. Given this focus on everyday functioning, Rabin et al.'s (2007) finding that ecologically valid measures are infrequently used in clinical practice is concerning.

One solution to the dilemma facing neuropsychologists may be to use Sbordone's (1996) vector analysis. This method was initially proposed to assess whether or not information obtained from neuropsychological measures was ecologically valid.

However, it can also be useful in determining the type of measures, verisimilitude or veridicality, to use with clients. Sbordone proposed that test data be compared with all collateral information about a client. If the neuropsychological results corroborate the collateral information, then the results of the measures may be viewed as being ecologically valid. On the other hand, if the information is contradictory, then the measures are not likely to be providing ecologically valid information. In these latter cases, incorporating recently developed measures, such as the one developed in the current series of studies, into the neuropsychological assessment will help neuropsychologists predict their client's functional ability. Once this is known, rehabilitation strategies can be developed to improve everyday functioning.

Limitations and Future Research

The cut-off points for the functional ratings were selected somewhat arbitrarily, as there was no external criterion available to help make the decision. It was assumed that if participants performed these steps without displaying executive difficulties, their performance suggested intact executive skills (i.e., *Full Functioning*), whereas if they committed at least one error, they received a score that suggests some level of executive difficulty. However, this assumption does not take into account the fact that it is common for people in the general population to display errors when performing activities of daily living (see Bottari, Swaine, & Dutil, 2007). Thus, a participant with TBI could be reported as displaying executive dysfunction, when in reality, they simply displayed a commonly observed error. Although the functional ratings from *Full Function With Some Limitations* to *Apparent Total Disability* allow for behavioural errors, as each rating is associated with a specific range of errors, the same is not true for the rating *Full Function*. Therefore, more flexibility in the cut-off point for achieving *Full Function* (e.g., between zero and two errors) may be warranted to allow for commonly performed errors.

One of the limitations of behavioural observations in general is that internal cognitive processes cannot be observed. If a participant purposely performs a behaviour that appears to be a cognitive activity limitation and does not notify the examiner, then the purposeful behaviour may be scored incorrectly. An example from the functional task highlights this methodological limitation. One of the participants in the normative group added twice as much baking powder to the mixture than was called for in the recipe. She was rated as having performed an Executive Memory error as she did not add the correct amount of baking powder. After the task was completed, and without prompting from the examiner, the participant reported why she added the extra baking powder. She explained that she does not bake very often and, as a result, the baking powder that she has at home is typically past the use-by-date. Therefore, she regularly adds twice as much baking powder as called for to compensate for her expired baking powder. Although the brownies were prepared with baking powder provided by the examiner, which was not expired, she said that she was unsure as to the age of the baking powder and added twice as much as a precautionary measure. One way of assessing internal, cognitive, processes would be to ask participants to comment on their behaviour after completing of the functional task. In this way, the examiner would be able to identify purposeful behaviours that would otherwise be recorded as evidence of executive dysfunction.

In the current series of studies, participants were not asked to comment on their behaviour during the functional task. It was thought that if participants were asked to do so, this may cue them to monitor their behaviour in a way that was different from their usual monitoring behaviour. Of course, the simple fact of observing participants may alter their performance such that it is different from their ‘usual‘ performance. This confounding variable is inherent to all behavioural observation measures in which participants are aware that they are being observed (see Chiesa & Hobbs, 2008).

The injury severity of participants with TBIs were not assessed in this study. However, this is mitigated somewhat by the injury severity estimate that was made—22 participants likely sustained a severe TBI. Furthermore, the validity of the study was assessed, in part, through a correlational analysis. Given that participants act as their own controls in this type of analysis, severity of injury is controlled for.

Another limitation to the study was that raters were not blind to the group status of participants. Although biases in rating can arise due to the representativeness and availability heuristics (see Tversky & Kahneman, 1974), Wolfson et al. (2000) found that raters were more influenced by the ratings of hypothetical team members than by demographic variables, such as group affiliation. In the Main Assessment, raters did not discuss the performance of any participant prior to scoring the functional task and the only time that participant performance was discussed was in the few instances when clarification was necessary to ensure that the second rater followed the scoring manual correctly. Furthermore, each participant performed the functional task first, followed by the formal neuropsychological measures. This step was taken to ensure that the raters were not influenced by the information obtained from the formal measures. In this way, raters were unaware of the executive ability of each participant, only that a particular participant had had a brain injury or not.

Despite the positive results of the study, every participant received an Initiation score of *Effective Function*. In other words, there were no participants who displayed pathological inertia (see Lezak et al., 2004). People with TBI are unlikely to display inertia, or any other type of executive dysfunction for that matter, if tests are highly structured (Bennett, 2001; Howieson & Lezak, 2004). In the functional task, participants were given a recipe and instructed to begin cooking. Therefore, although the task was more unstructured than formal neuropsychological measures, it may not have been unstructured enough to elicit initiation difficulties. Alternatively, the functional rating itself may not have been sensitive enough to detect difficulties in this area. The scoring manual stated that raters were required to distinguish initiation difficulties from planning and organisation difficulties. It may have been that initiation difficulties were identified, but were scored as planning and organisation errors. An example may help clarify this point: A participant places margarine in a bowl and melts it as per the recipe. They then attempt to add sugar to the bowl, but do not, due to pathological inertia. Despite their inertia, they are able to initiate the next step, which is adding eggs to the mixture. In this case, raters would likely score the participant as making a planning and organisation error, rather than an initiation error. This ‘incorrect’ scoring of participant behaviour would remain unless the participant indicated that they had intended to add the sugar, but were unable to due so due to initiation difficulties.

One potential criticism of the current study may be that Bonferroni adjustments were not made in the statistical analysis of the Main Assessment. Thus, the statistically significant correlations found between the functional task and the formal neuropsychological measures may simply be Type I errors. Despite this possibility, Bonferroni adjustments were not made for several reasons. First, the current study was an exploratory study designed to determine whether or not a suitable scoring method could be designed, rather than a validation study of an established measure. Second, statistically significant results were not expected. Therefore, ensuring that the statistically significant results were indeed significant was not necessary. Third, it was deemed that effect size was more important to a study of this nature that statistical significance, hence the use of Pearson's correlations. Finally, one of the problems with Bonferroni adjustments is that in controlling for Type I errors, there is an increase in the chances of making Type II errors (Perneger, 1998). This final reason led Nakagawa (2004) to discourage the use of Bonferroni adjustment as a matter of course in scientific research.

The results of the Main Assessment provide initial support for the ecological validity of the functional task. However, the functional rating Sustained and Directed Attention did not yield moderate correlations with the formal measures of EF. One reason for this may be that the formal measures used in the study do not assess this aspect of EF. Therefore, future research should incorporate a specific sustained and directed attention measure, such as the Elevator Counting or Lottery subtests from the Test of Everyday Attention (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994), to determine the ecological validity of this functional rating.

Finally, the functional task yielded significant and moderate, non-significant, correlations with formal measures of EF. Although this was interpreted as providing initial support for the ecological validity of the functional task, it may also be interpreted as simply providing support for the ecological validity of the Sorting and Tower Tests. Furthermore, the ecological validity of the functional task is further called into question based on the association this task had with the EBIQ. Although it was anticipated that positive correlations would be found between the functional task and the EBIQ, this was generally not the case. Thus, in order for the functional task to display ecological validity, performance on the task needs to be significantly related to performance in everyday life (i.e., on measures such as the EBIQ or Dysexecutive Questionnaire).

Overall Conclusion

Following TBI, people will likely experience impairments in various cognitive processes, with executive dysfunction being one of the most common (Stuss & Levine, 2002). A host of standardised measures have been created over the years to determine impairments in cognitive functioning. However, this information may not allow neuropsychologists to accurately predict a person's everyday functioning as the relationship between cognitive impairments and daily functioning is not a linear one. Rather, a number of factors influence a person's everyday functioning after sustaining a TBI. Using the ICF in case formulation cues clinicians to consider these various factors as functioning should be assessed in both structured and natural environments and with and without compensatory interventions (Reed et al., 2005). In a similar vein, Wilson (1996) stated that clinicians should incorporate real-world tasks into the formal neuropsychological assessment process. Although Wilson made this remark over a decade ago, few have heeded her call (see Rabin et al., 2007).

One reason for the current situation may be the relative absence of cognitive measures of everyday behaviours (Oddy et al., 1999). For this reason, the current study attempted to develop and trial a systematic behavioural observation approach in quantifying ongoing cognitive activity limitations that tapped EF after TBI. In order to determine a suitable "real-world" task to be used in the study, participants in the Pre-Study Survey were asked about their daily behaviours. Information obtained from this survey indicated that participants regularly engaged in cooking tasks. Furthermore, they also reported experiencing difficulty independently completing these tasks. Thus, a cooking task was chosen as the "real-world" task as it may be considered an ADL and is sufficiently complex to ensure that EF underlies a significant part of the successful completion of the task.

The cooking task developed in the Developmental Trial required participants to make a chocolate brownie and a hot drink. Two manuals were created to ensure consistency in task administration and scoring. Results from this trial indicated that the functional task displayed adequate reliability. The Main Assessment was then conducted to determine the validity of the task. Results from this assessment indicated that, in addition to displaying adequate inter-rater reliability, the functional task yielded significant and non-significant, moderate, correlations to formal neuropsychological measures of EF. In addition, significant correlations were also found between the functional task and self-reported cooking ability. Finally, the functional task could be used to differentiate participants who had sustained a TBI from participants in the normative group. Thus, although this study represents a preliminary trial of a new behavioural observation assessment approach, it shows promise as a potentially useful measure for assessing executive functioning.

References

Achenbauch, T., McConaughy, S., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

Adams, R., Smigielski, J., & Jenkins, R. (1984). Development of a Satz-Mogel short form of the WAIS-R. *Journal of Consulting and Clinical Psychology, 52*, 908.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9*, 357-381.

Alvarez, J., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*, 17-42.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th text revision ed.). Washington, DC: Author.

Angoff, W. H. (1984). *Scales, norms and equivalent scores*. Princeton, NJ: Educational Testing Services.

Arlinghaus, K., Shoaib, A., & Price, T. (2005). Neuropsychiatric assessment. In J. M. Silver, T. W. McAllister & S. C. Yudofsky (Eds.), *Textbook of traumatic brain injury* (pp. 59-78). Washington, DC: American Psychiatric Publishing, Inc.

Armitage, S. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychology Monographs, 60*, 1-48.

Baguena, N., Thomas-Anterion, C., Sciessere, K., Truche, A., Extier, C., Guyot, E., et al. (2006). Ecologic evaluation in the cognitive assessment of brain injury patients: generation and execution of script. *Annales de Readaptation et de Medecine Physique, 49*, 234-241.

Bamdad, M., Ryan, L., & Warden, D. (2003). Functional assessment of executive abilities following traumatic brain injury. *Brain Injury, 17*, 1011-1020.

Barbizet, J., & Duizabo, P. (1977). *Neuropsychologie*. Paris: Masson.

Baron, I. S. (2004). Delis-Kaplan executive function system. *Child Neuropsychology, 10*, 147-152.

Baum, C., & Edwards, D. F. (1993). Cognitive performance in senile dementia of the Alzheimer's type: The kitchen task assessment. *American Journal of Occupational Therapy, 47*, 431-436.

Bennett, P., Ong, B., & Ponsford, J. (2005). Measuring executive dysfunction in an acute rehabilitation setting: Using the dysexecutive questionnaire (DEX). *Journal of the International Neuropsychological Society, 11*, 376-385.

Bennett, T. (2001). Neuropsychological evaluation in rehabilitation planning and evaluation of functional skills. *Archives of Clinical Neuropsychology, 16*, 237-253.

Benton, A. (1968). Differential behavioural effects in frontal lobe disease. *Neuropsychologia, 6*, 53-60.

Bergner, M., Bobbitt, R., Pollard, W., Martin, D., & Gilson, B. (1976). The sickness impact profile: Validation of a health status measure. *Medical Care, 14*, 57-67.

Bernspång, B. (1995). Differences between persons with right or left cerebral vascular accident on the assessment of motor and process skills. *Archives of Physical Medicine & Rehabilitation, 76*, 114-1151.

Bettcher, B. M., Giovannetti, R., & MacMullen, L. (2008, February). *Correction of everyday action errors in dementia: Time frames and neuropsychological correlates*. Poster session presented at the Thirty-Sixth Annual International Neuropsychological Society conference, Waikoloa, Hawaii.

Bilbao, A., Kennedy, C., Chatterji, S., Üstün, B., Barquero, J. L. V., & Barth, J. (2003). The ICF: Applications of the WHO model of functioning, disability and health to brain injury rehabilitation. *Neurorehabilitation, 18*, 239-250.

Birkimer, J., & Brown, J. (1979). A graphical judgemental aid which summarizes obtained and change reliability data and helps assess the believability of experimental effects. *Journal of Applied Behavior Analysis, 12*, 523-533.

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *BMJ, 314*, 572.

Boone, K., Lesser, I. M., Hill-Gutierrez, E., Berman, N. G., & D'Elia, L. (1993). Rey-Osterrieth complex figure performance in healthy, older adults: Relationship to age, education, sex, and IQ. *Clinical Neuropsychologist, 7*, 22-28.

Boone, K., Miller, B., Lesser, I., Hill, E., & D'Elia, L. (1990). Performance on frontal lobe tests in healthy, older individuals. *Developmental Neuropsychology, 6*, 215-223.

Boone, K., Pontón, M., Gorsuch, R., González, J., & Miller, B. (1998). Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology, 13*, 585-595.

Bottari, C., Swaine, B., & Dutil, E. (2007). Interpreting activity of daily living errors for treatment and discharge planning: the perception of occupational therapists. *Journal of Head Trauma Rehabilitation, 22*, 26-30.

Boyd, T., & Sautter, S. (1993). Route-finding: A measure of everyday executive functioning in the head-injured adult. *Applied Cognitive Psychology, 7*, 171-181.

Braver, T., Barch, D., Gray, J., Molfese, D., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral Cortex, 11*, 825-836.

Brooks, D. (1990). Cognitive deficits. In M. Rosenthal, E. R. Griffith, M. R. Bond & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury* (2nd ed., pp. 163-178). Philadelphia, PA: F A Davis.

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

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

Burgess, P. W., Alderman, N., Evans, J., Wilson, B. A., Emslie, H., & Shallice, T. (1996). Modified six element test. In B. A. Wilson, N. Alderman, P. W. Burgess, H. Emslie & J. J. Evans (Eds.), *Behavioural assessment of the dysexecutive syndrome*. Dury St. Edmunds, U.K.: Thames Valley Test Company.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M., Dawson, D. R., et al. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society, 12*, 194-209.

Burgess, P. W., Alderman, N., Wilson, B. A., Evans, J., & Emslie, H. (1996). The dysexecutive questionnaire. In B. A. Wilson, N. Alderman, P. W. Burgess, H. Emslie & J. J. Evans (Eds.), *Behavioural assessment of the dysexecutive questionnaire*. Bury St. Edmunds, U.K.: Thames Valley Test Company.

Burgess, P. W., & Simons, J. S. (2005). Theories of frontal lobe executive function: clinical applications. In P. W. Halligan & D. T. Wade (Eds.), *Effectiveness of rehabilitation for cognitive deficits* (pp. 211-231). New York: Oxford University Press.

Bürk, K., Daum, I., & Rüb, U. (2006). Cognitive function in multiple system atrophy of the cerebellar type. *Movement Disorders, 21*, 772-776.

Busch, R., McBride, A., Curtiss, G., & Vanderploeg, R. (2005). The components of executive functioning in traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 27*, 1022-1032.

Cassidy, J. D., Carroll, L., Peloso, P., Borg, J., von Holst, H., Holm, L., et al. (2004). Incidence, risk factors and prevention of mild traumatic brain injury: Results of the WHO collaborating centre task force on mild traumatic brain injury. *Journal of Rehabilitation Medicine, 36*(Suppl. 43), 28-60.

Chamberlain, E. (2003). Behavioural assessment of the dysexecutive syndrome (BADS). *Journal of Occupational Psychology, Employment and Disability, 5*, 33-37.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review, 13*, 181-197.

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology, 21*, 217-227.

Chaytor, N., Temkin, N., Machamer, J., & Dikmen, S. (2007). The ecological validity of neuropsychological assessment and the role of depressive symptoms in moderate to severe traumatic brain injury. *Journal of the International Neuropsychological Society, 13*, 377-385.

Chevignard, M., Pillon, B., Pradat-Diehl, P., Taillefer, C., Rousseau, S., Le Bras, C., et al. (2000). An ecological approach to planning dysfunction: script execution. *Cortex, 36*, 649-669.

Chevignard, M., Servant, V., Mariller, A., Abada, G., Pradat-Diehl, P., & Laurent-Vannier, A. (2009). Assessment of executive functioning in children after TBI with a naturalistic open-ended task: A pilot study. *Developmental Neurorehabilitation, 12*, 76-91.

Chevignard, M., Taillefer, C., Picq, C., Poncet, F., Noulhiane, M., & Pradat-Diehl, P. (2008). Ecological assessment of the dysexecutive syndrome using execution of a cooking task. *Neuropsychological Rehabilitation, 18*, 461-485.

Chiesa, M., & Hobbs, S. (2008). Making sense of social research: How useful is the Hawthorne effect? *European Journal of Social Psychology, 38*, 67-74.

Code of Ethics Review Group. (2002). *Code of ethics for psychologists working in Aotearoa/New Zealand*. Wellington, New Zealand: New Zealand Psychologists Board, the New Zealand Psychological Society, and the New Zealand College of Clinical Psychologists.

Comalli, P. J., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test in childhood, adulthood, and aging. *Journal of Genetic Psychology, 100*, 47-53.

Constantinidou, F., Thomas, R. D., & Best, P. J. (2004). Principles of cognitive rehabilitation: An integrative approach. In M. J. Ashley (Ed.), *Traumatic brain injury: Rehabilitative treatment and case management* (2nd ed., pp. 337-365). Boca Raton, FL: CRC Press.

Crawford, J., & Henry, J. (2005). Assessment of executive dysfunction. In P. W. Halligan & D. T. Wade (Eds.), *Effectiveness of rehabilitation for cognitive deficits* (pp. 233-245). New York, NY: Oxford University Press.

Dath, D., Regehr, G., Birch, D., Schlachta, C., Poulin, E., Mamazza, J., et al. (2004). Toward reliable operative assessment: The reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surgical Endoscopy, 18*, 1800-1804.

Daum, I., Snitz, B. E., & Ackermann, H. (2001). Neuropsychological deficits in cerebellar syndromes. *International Review of Psychiatry, 13*, 268-275.

Deb, S., Bryant, E., Morris, P., Prior, L., Lewis, G., & Haque, S. (2007). Development and psychometric properties of the patient-head injury participation scale (P-HIPS) and the patient-head injury neurobehavioral assessment scale (P-HINAS): Patient and family determined outcomes scales. *Neuropsychiatric Disease and Treatment, 3*, 373-388.

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan executive function system*. San Antonio, TX: The Psychological Corporation.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (1987). *The California verbal learning test: Research edition*. San Antonio, TX: The Psychological Corporation.

Demakis, G. (2003). A meta-analytic review of the sensitivity of the Wisconsin card sorting test to frontal and lateralized frontal brain damage. *Neuropsychology, 17*, 255-264.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

DiLavore, P. C., Lord, C., & Rutter, M. (1995). The pre-linguistic autism diagnostic observation schedule. *Journal of Autism & Developmental Disorders, 25*, 355-379.

Dirette, D. K., & Plaisier, B. R. (2007). The development of self-awareness of deficits from 1 week to 1 year after traumatic brain injury: Preliminary findings. *Brain Injury, 21*, 1131-1136.

Downs, K. (n.d.). *Microwave brownie recipe*. Retrieved 23 April, 2007, from http://www.cdkitchen.com/recipes/recs/28/Microwave_Brownies43253.shtml

Dusart, I., Ghoumari, A., Wehrle, R., Morel, M., Bouslama-Oueghlani, L., Camand, E., et al. (2005). Cell death and axon regeneration of Purkinje cells after axotomy: Challenges of classical hypotheses of axon regeneration. *Brain Research - Brain Research Reviews, 49*, 300-316.

Elliott, R. (2003). Executive functions and their disorders. *British Medical Bulletin, 65*, 49-59.

Eranti, V., Janakiramaiah, N., Gangadhar, B., & Subbakrishna, D. (1998). Drug induced extrapyramidal symptoms: Rating using video-records. *NIMHANS Journal, 16*, 9-13.

Fisher, A. (2001a). *Assessment of motor and process skills. Vol. I: Development, standardization, and administration manual* (4th ed.). Fort Collins, CA: Three Star Press.

Fisher, A. (2001b). *Assessment of motor and process skills. Vol. 2: User Manual* (4th ed.). Fort Collins, CA: Three Star Press.

Flashman, L., Amador, X., & McAllister, T. (2005). Awareness of deficits. In J. M. Silver, T. W. McAllister & S. C. Yudofsky (Eds.), *Textbook of traumatic brain injury* (pp. 353-367). Washington, DC: American Psychiatric Publishing, Inc.

Fricke, J., Unsworth, C., & Worrell, D. (1993). Reliability of the functional independence measure with occupational therapists. *Australian Occupational Therapy Journal, 40*, 7-15.

Fürst, E. (1997). Cooking and femininity. *Women's Studies International Forum, 20*, 441-449.

Gennarelli, T., & Graham, D. (2005). Neuropathology. In J. M. Silver, T. W. McAllister & S. C. Yudofsky (Eds.), *Textbook of traumatic brain injury* (pp. 27-50). Washington, DC: American Psychiatric Publishing, Inc.

Gioia, D., & Brekke, J. S. (2009). Neurocognition, ecological validity, and daily living in the community for individuals with schizophrenia: A mixed methods study. *Psychiatry, 72*, 94-107.

Gioia, G., & Isquith, P. (2004). Ecological assessment of executive function in traumatic brain injury. *Developmental Neuropsychology, 25*, 135-158.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function*. Odessa, FL: Psychological Assessment Resources.

Gioia, G. A., Isquith, P. K., & Kenealy, L. E. (2008). Assessment of behavioral aspects of executive function. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A lifespan perspective* (pp. 179-202). Philadelphia, PA: Taylor & Francis.

Goldman-Rakic, P. (1984). The frontal lobes: Uncharted provinces of the brain. *Trends in Neurosciences, 7*, 425-429.

Graham, D. (1995). Neuropathology of head injury. In P. D. Nussbaum (Ed.), *Handbook of neuropathology and aging* (pp. 43-59). New York: McGraw Hill.

Graham, J. R., & Naglieri, J. A. (2003). *Handbook of psychology: Assessment psychology* (Vol. 10). Hoboken, NJ: John Wiley & Sons Inc.

Groswasser, Z., Cohen, M., & Keren, O. (1998). Female TBI patients recover better than males. *Brain Injury, 12*, 805-808.

Hagg, T. (2006). Collateral sprouting as a target for improved function after spinal cord injury. *Journal of Neurotrauma, 23*, 281-294.

Hall, K., Bushnik, T., Lakisic-Kazazic, B., Wright, J., & Cantagallo, A. (2001). Assessing traumatic brain injury outcome measures for long-term follow-up of community-based individuals. *Archives of Physical Medicine & Rehabilitation, 82*, 367-374.

Hall, K., Hamilton, B., Gordon, W., & Zasler, N. (1993). Characteristics and comparisons of functional assessment indices: Disability rating scale, functional independence measure, and functional assessment measure. *Journal of Head Trauma Rehabilitation, 8*, 60-74.

Halliday, A. (1999). Pathophysiology. In D. W. Marion (Ed.), *Traumatic brain injury* (pp. 29-38). New York: Thieme.

Hamilton, B., Granger, C., Sherwin, F., Zielezny, M., & Tashman, J. (1987). A uniform national data system for medical rehabilitation. In J. M. Fuhrer (Ed.), *Rehabilition outcomes: Analysis and measurement* (pp. 137-147). Baltimore, MD: Brookes.

Haneline, M. (2007). *Evidence-based chiropractic practice*. Sudbury, MA: Jones and Bartlett Publishers, Inc.

Hart, T., Whyte, J., Kim, J., & Vaccaro, M. (2005). Executive function and self-awareness of "real-world" behavior and attention deficits following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 20*, 333-347.

Haxby, J., Petit, L., Ungerleider, L., & Courtney, S. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *Neuroimage, 11*, 380-391.

Hayes, J., & Hatch, J. (1999). Issues in measuring reliability. *Written Communication, 16*, 354-367.

Heaton, R. (1981). *Wisconsin card sorting test manual*. Odessa, FL: Psychological Assessment Resources.

Hermansson, L. M., Bodin, L., & Eliasson, A.-C. (2006). Intra- and inter-rater reliability of the assessment of capacity for myoelectric control. *Journal of Rehabilitation Medicine, 38*, 118-123.

Heyder, K., Suchan, B., & Daum, I. (2004). Cortico-subcortical contributions to executive control. *Acta Psychologica, 115*, 271-289.

Holm, S., Schönberger, M., Poulsen, I., & Caetano, C. (2009). Patients' and relatives' experience of difficulties following severe traumatic brain injury: The sub-acute stage. *Neuropsychological Rehabilitation, 19*, 444-460.

Holmes-Bernstein, J., & Waber, D. P. (1990). Developmental neuropsychological assessment: The systemic approach. In A. A. Boulton, G. B. Baker & M. Hiscock (Eds.), *Neuropsychology* (pp. 311-371). Totowa, NJ: Humana Press.

Homack, S., Lee, D., & Riccio, C. (2005). Test review: Delis-Kaplan executive function system. *Journal of Clinical and Experimental Neuropsychology, 27*, 599-609.

Hoskin, K., Jackson, M., & Crowe, S. (2005). Can neuropsychological assessment predict capacity to manage personal finances? A comparison between brain impaired individuals with and without administrators. *Psychiatry, Psychology and Law, 12*, 56-67.

Howieson, D. B., & Lezak, M. D. (2004). The neuropsychological evaluation. In S. C. Yudofsky & R. E. Hales (Eds.), *Essentials of neuropsychiatry and clinical neurosciences* (pp. 55-80). Arlington, VA: American Psychiatric Publishing, Inc.

Josman, N., & Birnboim, S. (2001). Measuring kitchen performance: What assessment should we choose? *Scandinavian Journal of Occupational Therapy, 8*, 193-202.

Katz, D., & Alexander, M. (1994). Traumatic brain injury. Predicting course of recovery and outcome for patients admitted to rehabilitation. *Archives of Neurology, 51*, 661-670.

Kim, S. G., Uğurbil, K., & Strick, P. L. (1994). Activation of a cerebellar output nucleus during cognitive processing. *Science, 265*, 949-951.

Knight, C., Alderman, N., & Burgess, P. W. (2002). Development of a simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation, 12*, 231-256.

Kolb, B. (2004). Mechanisms of cortical plasticity after neuronal injury. In J. Ponsford (Ed.), *Cognitive and behavioral rehabilitation: From neurobiology to clinical practice* (pp. 30-58). New York, NY: Guilford Press.

Kolb, B., & Cioe, J. (2004). Neuronal organization and change after neuronal injury. In J. Ponsford (Ed.), *Cognitive and behavioral rehabilitation: From neurobiology to clinical practice* (pp. 7-29). New York: Guilford Press.

Kottorp, A., Bernspång, B., & Fisher, A. (2003). Validity of a performance assessment of activities of daily living for people with developmental disabilities. *Journal of Intellectual Disability Research, 47*, 597-605.

Landa-Gonzalez, B. (2001). Multicontextual occupational therapy intervention: A case study of traumatic brain injury. *Occupational Therapy International, 8*, 49-62.

Lewis, M., & Babbage, D. (2007a, April). Cognitive disability research. *Brainlink,* p. 5.

Lewis, M., & Babbage, D. (2007b, September). Functional ability research. *The Brain Injury Association of New Zealand, Wellington Incorporated,* p. 1.

Lewis, M., & Gibson, B. (2009). *Researchers launch major brain injury study.* Retrieved February 24, 2009, from http://www.massey.ac.nz/massey/about-us/news/article.cfm?mnarticle=researchers-launch-major-brain-injury-study-23-02-2009

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Lindén, A., Boschian, K., Eker, C., Schalén, W., & Nordström, C.-H. (2005). Assessment of motor and process skills reflects brain-injured patients' ability to resume independent living better than neuropsychological tests. *Acta Neurologica Scandinavica, 111*, 48-53.

Long, C. J. (1996). Neuropsychological tests: A look at our past and the impact that ecological issues may have on our future. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 1-14). Delray Beach, FL: GR Press/St Lucie Press.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism & Developmental Disorders, 30*, 205-223.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism & Developmental Disorders, 19*, 185-212.

Lucas, J. (1998). Traumatic brain injury and postconcussive syndrome. In P. J. Snyder & P. D. Nussbaum (Eds.), *Clinical neuropsychology: A pocket handbook for assessment* (pp. 243-265). Washington, DC: American Psychological Association.

MacDonald, A., Cohen, J., Stenger, V., & Carter, C. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835-1838.

Mahoney, F., & Barthel, D. (1965). Functional evaluation: The Barthel index. *Maryland State Medical Journal, 14*, 61-65.

Malec, J., Ivnik, R., & Hinkeldey, N. (1991). Visual Spatial Learning Test. *Psychological Assessment, 3*, 82-88.

Malec, J., Moessner, A., Kragness, M., & Lezak, M. (2000). Refining a measure of brain injury sequelae to predict postacute rehabilitation outcome: Rating scale analysis of the Mayo-Portland adaptability inventory. *Journal of Head Trauma Rehabilitation, 15*, 670-682.

Malloy, P., Cohen, R., & Jenkins, M. (1998). Frontal lobe function and dysfunction. In P. J. Snyder & P. D. Nussbaum (Eds.), *Clinical neuropsychology: A pocket handbook for assessment* (pp. 573-590). Washington, DC: American Psychological Association.

Manchester, D., Priestley, N., & Jackson, H. (2004). The assessment of executive functions: Coming out of the office. *Brain Injury, 18*, 1067-1081.

McCue, M., & Pramuka, M. (1998). Functional assessment. In G. Goldstein & S. R. Beers (Eds.), *Rehabilitation* (pp. 113-129). New York, NY: Plenum Press.

McCullagh, S., & Feinstein, A. (2005). Cognitive changes. In J. M. Silver, T. W. McAllister & S. C. Yudofsky (Eds.), *Textbook of traumatic brain injury* (pp. 321-335). Washington, DC: American Psychiatric Publishing, Inc.

McDonald, B., Flashman, L., & Saykin, A. (2002). Executive dysfunction following traumatic brain injury: Neural substrates and treatment strategies. *Neurorehabilitation, 17*, 333-344.

McKenna, P., Mortimer, A., & Hodges, J. (1994). Semantic memory and schizophrenia. In A. S. David & J. C. Cutting (Eds.), *The neuropsychology of schizophrenia*. Hove, U.K.: Erlbaum.

McPheeters, H. (1984). Statewide mental health outcome evaluation: A perspective of two southern states. *Community Mental Health Journal, 20*, 44-55.

Miller, E. (1984). Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of pathology. *British Journal of Clinical Psychology, 23*, 53-57.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9*, 100-110.

Mitchell, M., & Miller, L. (2008). Prediction of functional status in older adults: The ecological validity of four Delis-Kaplan Executive Function System tests. *Journal of Clinical and Experimental Neuropsychology, 30*, 683-690.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience, 21*, 7733-7741.

Morse, P., & Montgomery, C. (1992). Neuropsychological evaluation of traumatic brain injury. In R. F. White (Ed.), *Clinical syndromes in adult neuropsychology: The practitioner's handbook* (pp. 85-176). New York, NY: Elsevier Science.

Nakagawa, S. (2004). A farwell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology, 15*, 1044-1045.

Neistadt, M. E. (1992). The Rabideau kitchen evaluation - revised: An assessment of meal preparation skills. *The Occupational Therapy Journal of Research, 12*, 242-255.

Neistadt, M. E. (1994). A meal preparation treatment protocol for adults with brain injury. *American Journal of Occupational Therapy, 48*, 431-438.

Nelson, H. (1976). A modified card sorting test sensitive to frontal lobe damage. *Cortex, 12*, 323-324.

Nelson, H., & Willison, J. (1991). *National adult reading test (NART)*. Windsor, U.K.: NFER-Nelson.

New Zealand Guidelines Group. (2006). *Evidence-based best practice guidelines. Traumatic brain injury: Diagnosis, acute management and rehabilitation*. Wellington, New Zealand: Accident Compensation Corporation.

Newton, K. (2009, March 16). New way of assessing brain injury. *The Dominion Post,* p. A9.

Norris, G., & Tate, R. L. (2000). The behavioural assessment of the dysexecutive syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation, 10*, 33-45.

Oddy, M., Alcott, D., Francis, E., Jenkins, K., & Fowlie, C. (1999). Methods of evaluation in a cognitive-behavioural rehabilitation programme for brain injury: The experience of Ticehurst House and Unsted Park Hospitals. *Neuropsychological Rehabilitation, 9*(3-4), 373-384.

Odhuba, R., van den Broek, M., & Johns, L. (2005). Ecological validity of measures of executive functioning. *British Journal of Clinical Psychology, 44*, 269-278.

Olver, J., Ponsford, J., & Curran, C. (1996). Outcome following traumatic brain injury: A comparison between 2 and 5 years after injury. *Brain Injury, 10*, 841-848.

Ottochian, M., Salim, A., Berry, C., Chan, L. S., Wilson, M. T., & Margulies, D. R. (2009). Severe traumatic brain injury: Is there a gender difference in mortality? *The American Journal of Surgery, 197*, 155-158.

Pagulayan, K. F., Temkin, N. R., Machamer, J. E., & Dikmen, S. S. (2007). The measurement and magnitude of awareness difficulties after traumatic brain injury: A longitudinal study. *Journal of the International Neuropsychological Society, 13*, 561-570.

Parizel, P. M., Ozsarlak, O., Van Goethem, J. W., van den Hauwe, L., Dillen, C., Verlooy, J., et al. (1998). Imaging findings in diffuse axonal injury after closed head trauma. *European Radiology, 8*, 960-965.

Perneger, T. (1998). What's wrong with Bonferroni adjustments. *BMJ, 316*, 1236-1238.

Peterson, D. (2005). International classification of functioning, disability and health: An introduction for rehabilitation psychologists. *Rehabilitation Psychology, 50*, 105-112.

Piotrowski, C., & Keller, J. W. (1989). Psychological testing in outpatient mental health facilities: A national study. *Professional Psychology: Research and Practice, 20*, 423-425.

Ponsford, J., Harrington, H., Olver, J., & Roper, M. (2006). Evaluation of a community-based model of rehabilitation following traumatic brain injury. *Neuropsychological Rehabilitation, 16*, 315-328.

Povlishock, J., & Katz, D. (2005). Update of neuropathology and neurological recovery after traumatic brain injury. *Journal of Head Trauma Rehabilitation, 20*, 76-94.

Prigatano, G., Altman, I., & O'Brien, K. (1990). Behavioral limitations that brain injured patients tend to underestimate. *Clinical Neuropsychologist, 4*, 163-176.

Quinsey, K., Findlay, C., & Willmott, L. (2005). *FIM information and procedures manual*. Wollongong, New South Wales: University of Wollongong, Australasian Rehabilitation Outcome Centre.

Rabin, L., Burton, L., & Barr, W. (2007). Utilization rates of ecologically oriented instruments among clinical neuropsychologists. *Clinical Neuropsychologist, 21*, 727-743.

Ready, R., Stierman, L., & Paulsen, J. (2001). Ecological validity of neuropsychological and personality measures of executive functions. *Clinical Neuropsychologist, 15*, 314-323.

Reed, G., Lux, J., Bufka, L., Trask, C., Peterson, D., Stark, S., et al. (2005). Operationalizing the international classification of functioning, disability and health in clinical settings. *Rehabilitation Psychology, 50*, 122-131.

Researchers launch major brain injury study. (2009, April). *Brainlink, 41,* 5.

Riva, D., & Giorgi, C. (2000). The cerebellum contributes to higher functions during development: Evidence from a series of children surgically treated for posterior fossa tumours. *Brain, 123*, 1051-1061.

Robertson, I., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The test of everyday attention*. Bury St. Edmunds, UK: Thames Valley Test Company.

Sbordone, R. J. (1996). Ecological validity: Some critical issues for the neuropsychologist. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 15-41). Delray Beach, FL: GR Press/St. Lucie Press.

Sbordone, R. J., & Guilmette, T. J. (1999). Ecological validity: Prediction of everyday and vocational functioning from neuropsychological test data. In J. J. Sweet (Ed.), *Forensic neuropsychology: Fundamentals and practice* (pp. 227-254). Lisse, The Netherlands: Swets and Zeitlinger.

Scheff, S., Price, D., Hicks, R., Baldwin, S., Robinson, S., & Brackney, C. (2005). Synaptogenesis in the hippocampal CA1 field following traumatic brain injury. *Journal of Neurotrauma, 22*, 719-732.

Schmahmann, J. D., & Sherman, J. C. (1997). Cerebellar cognitive affective syndrome. *International Review of Neurobiology, 41*, 433-440.

Schwartz, M. F., Segal, M., Veramonti, T., Ferraro, M., & Buxbaum, L. J. (2002). The naturalistic action test: A standardised assessment for everyday action impairment. *Neuropsychological Rehabilitation, 12*, 311-339.

Semkovska, M., Bedard, M.-A., Godbout, L., Limoge, F., & Stip, E. (2004). Assessment of executive dysfunction during activities of daily living in schizophrenia. *Schizophrenia Research, 69*, 289-300.

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

Shallice, T., & Evans, M. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex, 14*, 294-303.

Shaw, N. (2002). The neurophysiology of concussion. *Progress in Neurobiology, 67*, 281-344.

Sheskin, D. J. (2004). *Handbook of parametric and nonparametric statistical procedures* (3rd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Skinner, A., & Turner-Stokes, L. (2006). The use of standardized outcome measures in rehabilitation centres in the UK. *Clinical Rehabilitation, 20*, 609-615.

Smith, D., Meaney, D., & Shull, W. (2003). Diffuse axonal injury in head trauma. *Journal of Head Trauma Rehabilitation, 18*, 307-316.

Sohlberg, M., & Mateer, C. (2001). *Cognitive rehabilitation: An integrative neuropsychological approach*. New York: The Guilford Press.

Spence, M., Moss-Morris, R., & Chalder, T. (2005). The behavioural responses to illness questionnaire (BRIQ): A new predictive measure of medically unexplained symptoms following acute infection. *Psychological Medicine, 35*, 583-593.

Stein, D. G. (2001). Brain damage, sex hormones and recovery: A new role for progesterone and estrogen? *Trends in Neurosciences, 24*, 386-391.

Stocchi, F., & Brusa, L. (2000). Cognition and emotion in different stages and subtypes of Parkinson's disease. *Journal of Neurology, 247*(Suppl 2), II114-121.

Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.

Stuss, D., & Alexander, M. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research, 63*, 289-298.

Stuss, D., & Levine, B. (2002). Adult clinical neuropsychology: Lessons from studies of the frontal lobes. *Annual Review of Psychology, 53*, 401-433.

Svendsen, H., Teasdale, T., & Pinner, M. (2004). Subjective experience in patients with brain injury and their close relatives before and after a rehabilitation programme. *Neuropsychological Rehabilitation, 14*, 495-515.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet, 2*(7872), 81-84.

Teasdale, G., Murray, G., Parker, L., & Jennett, B. (1979). Adding up the Glasgow coma score. *Acta Neurochirurgica - Supplementum, 28*, 13-16.

Teasdale, T., Christensen, A., Willmes, K., Deloche, G., Braga, L., Stachowiak, F., et al. (1997). Subjective experience in brain-injured patients and their close relatives: A European brain injury questionnaire study. *Brain Injury, 11*, 543-563.

Tepper, S., Beatty, P., & DeJong, G. (1996). Outcomes in traumatic brain injury: Self-report versus report of significant others. *Brain Injury, 10*, 575-581.

Turner-Stokes, L. (2002). Standardized outcome assessment in brain injury rehabilitation for younger adults. *Disability and Rehabilitation: An International Multidisciplinary Journal, 24*, 383-389.

Turner-Stokes, L., Nyein, K., Turner-Stokes, T., & Gatehouse, C. (1999). The UK FIM +FAM: Development and evaluation. *Clinical Rehabilitation, 13*, 277-287.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

van den Broek, M., Bradshaw, C., & Szabadi, E. (1993). Utility of the modified Wisconsin card sorting test in neuropsychological assessment. *British Journal of Clinical Psychology, 32*, 333-343.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2008). A large-scale cross-sectional and longitudinal study into the ecological validity of neuropsychological test measures in neurologically intact people. *Archives of Clinical Neuropsychology, 23*, 787-800.

Van der Linden, M., Coyette, F., & Seron, X. (1992). Selective impairment of the "central executive" component of working memory: A single case study. *Cognitive Neuropsychology, 9*, 301-326.

Wechsler, D. (1981). *The Wechsler adult intelligence scale - revised.* New York: The Psychological Corporation.

Whiting, W., & Zernicke, R. (1998). *Biomechanics of musculoskeletal injury*. Champaign, IL: Human Kinetics.

Wilson, B. A. (1996). The ecological validity of neuropsychological assessment after severe brain injury. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 413-428). Delray Beach, FL: GR Press/St Lucie Press.

Wilson, B. A. (2000). Compensating for cognitive deficits following brain injury. *Neuropsychology Review, 10*, 233-243.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome*. Bury St. Edmunds, England: Thames Valley Test Company.

Wolfson, A., Doctor, J., & Burns, S. (2000). Clinician judgments of functional outcomes: How bias and perceived accuracy affect rating. *Archives of Physical Medicine & Rehabilitation, 81*, 1567-1574.

Wood, R. L., & Liossi, C. (2006). The ecological validity of executive tests in a severely brain injured sample. *Archives of Clinical Neuropsychology, 21*(5), 429-437.

World Health Organization. (1992-1994). *International statistical classification of disease and related health problems* (10th Rev. ed.). Geneva: Author.

World Health Organization. (2001). *The international classification of functioning, disability and health (ICF)*. Geneva: Author.

Wrightson, P., & Gronwall, D. (1998). Mild head injury in New Zealand: Incidence of injury and persisting symptoms. *New Zealand Medical Journal, 111*(1062), 99-101.

Appendices

Appendix A

Activities of Daily Life Questionnaire

Rate your current ability to perform these activities. Alternatively, rate your family member's current ability to perform these activities. Also, note the number of times the activity occurred during the past week.

Can not complete = Can not complete activity. Completely dependent on others.
Do not complete = Can complete activity, but someone else always completes the activity for me.
A lot of help needed = Can complete activity, but need a lot of help from others
A little help needed = Can complete activity, but need a little help from others
Independent = Can complete activity independently
N/A = Not applicable. Do not need to complete activity.

		Circle one answer						Circle the group that applies
1	Make a simple meal, that requires no cooking (e.g., a sandwich).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred 0 1-2 3-5 6-9 10-30 times last week.
2	Make a complex meal, that requires cooking (e.g., Spaghetti Bolognese).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
3	Make a cup of coffee or tea.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
4	Do housework (e.g., vacuuming, cleaning, and laundry).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
5	Post a letter at the Post Office.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
6	Go to the bank.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30

								times last week.
7	Pay the bills.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
8	Manage your finances (e.g., balance chequebook or moving money between different bank accounts).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
9	Organise daily routine (e.g., set-up appointments and activities)	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
10	Use the Yellow or White Pages to look up a phone number.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
11	Maintain personal hygiene (e.g., wash body, brush teeth).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
12	Use the internet.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
13	Use public transportation.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
14	Drive a vehicle.	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.
15	Go grocery shopping (e.g., selecting and buying food).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.

16	Separate recycling (e.g., paper versus plastics).	Can not complete	Do not complete	A lot of help needed	A little help needed	Independent	N/A	The activity occurred: 0 1-2 3-5 6-9 10-30 times last week.

17 Please list any daily activities that you, or your family member, have difficulty performing that were not included in this questionnaire.

Please list any strategies used to help complete the activities listed in the questionnaire.

	Topic	Rehabilitation Strategy
1	Simple meal	
2	Complex meal	
3	Coffee or tea	
4	Housework	
5	Post Office	

6 Bank

7 Pay bills

8 Finances

9 Daily routine

10 Yellow Pages

11 Hygiene

12 Internet

13 Public transport

14 Drive vehicle

15 Groceries

____

16 Recycling

____

17 ____

____

____

Activities of Daily Life Questionnaire

Please respond to each question with one of the answers below. Also, how many times did the activity occur during the past week.

Can not complete = Can not complete activity. Completely dependent on others.

Do not complete = Can complete activity, but someone else always completes the activity for me.

A lot of help needed = Can complete activity, but need a lot of help from others.

A little help needed = Can complete activity, but need a little help from others.

Independent = Can complete activity independently.

N/A = Not applicable. Do not need to complete activity.

The activity occurred 0 1-2 3-5 6-9 10-30 times last week.

Appendix B

EBIQ (European Brain Injury Questionnaire) – Self-Rating

Participant Identification: ______________________________

Date: ______________________________

This questionnaire is concerned with a number of problems or difficulties that people sometimes experience in their lives. We would like to know how much you have experienced any of these **within the last month**. Please listen to each item in the questionnaire and respond by indicating your answer ‘Not at all’ or ‘A little’ or ‘A lot’. Do not spend too much time on any item. Just give your most immediate response.

How much have you experienced the following?

1 Headaches ……………………………………………… Not at all A little A lot

2 Failing to get things done on time …………………..… Not at all A little A lot

3 Reacting too quickly to what others say or do ………..… Not at all A little A lot

4 Trouble remembering things …………………………… Not at all A little A lot

5 Difficulty participating in conversations ………………. Not at all A little A lot

6 Others do not understand your problems ……………… Not at all A little A lot

7 Everything is an effort …………………………………. Not at all A little A lot

8 Being unable to plan activities ………………………… Not at all A little A lot

9 Feeling hopeless about your future …………………….. Not at all A little A lot

10 Having temper outbursts ……………………………….. Not at all A little A lot

11 Being confused ………………………………………… Not at all A little A lot

12 Feeling lonely, even when together with other people … Not at all A little A lot

13 Mood swings without reason …………………………... Not at all A little A lot

14 Feeling critical of others ……………………………….. Not at all A little A lot

15 Having to do things slowly in order to be correct ……… Not at all A little A lot

16 Faintness or dizziness ………………………………….. Not at all A little A lot

17 Hiding your feelings from others ………………………Not at all A little A lot

18 Feeling sad ……………..……………………………Not at all A little A lot

19 Being ‘bossy’ or dominating ……………………………Not at all A little A lot

20 Needing to be reminded about personal hygiene ………Not at all A little A lot

21 Difficulty managing your finances …………………….. Not at all A little A lot

22 Trouble concentrating ………………………………….. Not at all A little A lot

23 Failing to notice other people’s moods ..………………Not at all A little A lot

24 Feeling anger against other people …………………….. Not at all A little A lot

25 Having your feelings easily hurt ………………………..Not at all A little A lot

26 Feeling unable to get things done ……………………… Not at all A little A lot

27 Annoyance or irritation …………………………………Not at all A little A lot

28 Problems with household chores ……………………… Not at all A little A lot

29 Lack of interest in hobbies at home ……………………Not at all A little A lot

30 Feeling lonely …………………………………………. Not at all A little A lot

31 Feeling inferior to other people …………………………Not at all A little A lot

32 Sleep problems …………………………………………Not at all A little A lot

33 Feeling uncomfortable in crowds ……………………… Not at all A little A lot

34 Shouting at people in anger ……………………………. Not at all A little A lot

35 Difficulty in communicating what you want to say ….... Not at all A little A lot

36 Being unsure what to do in dangerous situations ……… Not at all A little A lot

37 Being obstinate (stubborn) …………………………….. Not at all A little A lot

38 Lack of interest in your surroundings …………………. Not at all A little A lot

39 Thinking only of yourself ……………………………… Not at all A little A lot

40 Mistrusting other people ……………………………….. Not at all A little A lot

41 Crying easily ……………………………………………Not at all A little A lot

42 Difficulty finding your way in new surroundings …….... Not at all A little A lot

43 Being inclined to eat too much ………………………… Not at all A little A lot

44 Getting into quarrels easily ……………………………..Not at all A little A lot

45 Lack of energy or being slowed down …………………Not at all A little A lot

46 Forgetting the day of the week ………………………… Not at all A little A lot

47 Feeling of worthlessness ………………………………. Not at all A little A lot

48 Lack of interest in hobbies outside the home ………….. Not at all A little A lot

49 Needing help with personal hygiene ………………….. Not at all A little A lot

50 Restlessness …………………………………………. Not at all A little A lot

51 Feeling tense ………………………………………… Not at all A little A lot

52 Acting inappropriately in dangerous situations ……….. Not at all A little A lot

53 Feeling life is not worth living ……………………… Not at all A little A lot

54 Forgetting appointments ……………………………….. Not at all A little A lot

55 Leaving others to take the initiative in conversations …. Not at all A little A lot

56 Loss of sexual interest or pleasure …………………….. Not at all A little A lot

57 Throwing things in anger ……………………………… Not at all A little A lot

58 Preferring to be alone ………………………………….. Not at all A little A lot

59 Difficulty in making decisions ………………………… Not at all A little A lot

60 Losing contact with your friends ………………………. Not at all A little A lot

61 Lack of interest in current affairs ……………………… Not at all A little A lot

62 Behaving tactlessly …………………………………..... Not at all A little A lot

63 Having problems in general …………………………… Not at all A little A lot

Any other comment?

EBIQ (European Brain Injury Questionnaire)

Please respond to each question with one of the answers below.

Not At All   A Little   A Lot

Appendix C

Demographic Questionnaire

1. When were you born?

day	month	year
(e.g., 15)	(e.g., Jun)	(e.g., 1966)
☐☐	☐☐☐	☐☐☐☐

2. When did you sustain your brain injury?

day	month	year
(e.g., 15)	(e.g., Jun)	(e.g., 1966)
☐☐	☐☐☐	☐☐☐☐

3. Which ethnic group do you belong to?
(tick the circle or circles which apply)

○ New Zealand European
○ Māori
○ Samoan
○ Cook Island Maori
○ Tongan
○ Nuiean
○ Chinese
○ Indian
○ Other. Please state: ______________________

4. Who lives at home with you?
(tick the circle or circles which apply)

○ My legal husband or wife
○ My partner or de facto, boyfriend or girlfriend
○ My flatmates
○ Other. Please state: ______________________
○ None of the above. I live alone.

5. What is your highest academic qualification? ______________________________________

6. What was your main occupation at the time of your brain injury? ______________________

7. What is your current main occupation? ____

Appendix D

Information Sheets

Cognitive Disability after Brain Injury

Survey Information Sheet

You are invited to take part in a project to find out what kinds of activities people with brain injury perform on a daily basis and what kind of rehabilitation strategies they use to help them complete those activities.

What does taking part involve?

If you chose to take part, we will meet for a total of about thirty minutes. This meeting can take place either at the Psychology Clinic at Massey University in Wellington or at a place more convenient to you, but that is also suitable for the purposes of the project. This might include being able to complete the research at your home.

What will people in the study be asked to do?

The purpose of this research is find out what kinds of activities people with brain injury perform on a daily basis and what kinds of rehabilitation strategies they use to help them complete those activities. You will be asked a series of questions about the general activities you perform in an average day (e.g., cooking breakfast) and what kind of rehabilitation strategies (e.g., breaking tasks up into manageable parts) are used to help complete those activities. The questions also cover topics related to social, physical, and emotional functioning in the past month.

What information will researchers collect?

We will collect information by noting what you say during our visit.

Are there any benefits or risks?

The study does not have any direct benefits for you and is not designed to improve your rehabilitation. The procedures used in the study have no known expected risks or harmful effects.

What sorts of people are taking part in this study?

You may have been invited to participate in the study because you have been a client of the Psychology Clinic at Massey University or Cavit ABI Rehabilitation. If so, a clinician indicated you could be a suitable person to participate in this study. Alternatively, you may have been contacted through the Brain Injury Association. Volunteers in this study need to have had an acquired brain injury (e.g., traumatic brain injury, stroke) after the age of 16, have completed their acute rehabilitation following their brain injury, be comfortable using the English language, and have no significant visual impairment.

Taking part in this study is voluntary.

It is your choice whether you take part or not. You do not have to take part in this study. Also, if you do agree to take part, you can stop taking part at any stage without having to give a reason. Whatever you decide to do, the health care you receive will not be affected.

Please feel free to talk with a friend or family member if you need help to make the decision about whether to take part or not. Also, you are welcome to have a support person to be with you when we meet.

What will happen to the information collected in the study?

The things you say and the information we gather about you will be kept confidential and used for research purposes only. No material that identifies you will be used in any report on this project. We will store your information in a secure location, and only those involved in the project will be able to see it. We will store your information for seven years after the end of the project, after which time our records will be destroyed.

Finding out about the results of the study.

If you would like to find out the results of the project, please tick the first YES box on the consent form. After the project is completed, we will mail you the results.

We will be conducting a further study examining people's practical abilities to do some of these tasks in the future. We would like to invite all current participants to be involved in the follow up study. If you are happy for us to contact you again to discuss this, please tick the second YES box on the consent form.

Further information.

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington.

If you have any queries or concerns regarding your rights as a participant in this study you may wish to contact a Health and Disability Advocate, telephone:

- Auckland region 0800 555 050
- Wellington region 0800 42 36 38 (0800 4 ADNET)

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Mark Lewis	Duncan Babbage, PhD	Janet Leathem, PhD
<u>Doctoral Candidate</u>	<u>Clinic Director</u>	<u>Professor of Neuropsychology</u>

Cognitive Disability after Brain Injury

Family Member Information Sheet

You are invited to take part in a project to find out what kinds of activities people with brain injury perform on a daily basis and what kind of rehabilitation strategies they use to help them complete those activities.

What does taking part involve?

If you chose to take part, we will meet for a total of about thirty minutes. This meeting can take place either at the Psychology Clinic at Massey University in Wellington or at a place more convenient to you, but that is also suitable for the purposes of the project. This might include being able to complete the research at your home.

What will people in the study be asked to do?

The purpose of this research is find out what kinds of activities people with brain injury perform on a daily basis and what kinds of rehabilitation strategies they use to help them complete those activities. You will be asked a series of questions about the general activities your family member performs in an average day (e.g., cooking breakfast) and what kind of rehabilitation strategies (e.g., breaking tasks up into manageable parts) are used to help complete those activities. The questions also cover topics related to the social, physical, and emotional functioning of your family member in the past month.

What information will researchers collect?

We will collect information through the use of two questionnaires.

Are there any benefits or risks?

The study does not have any direct benefits for your family member and is not designed to improve their rehabilitation. The procedures used in the study have no known expected risks or harmful effects.

What sorts of people are taking part in this study?

You may have been invited to participate in the study because your family member has been a client of the Psychology Clinic at Massey University or Cavit ABI Rehabilitation. If so, a clinician indicated you could be a suitable person to participate in this study. Alternatively, you may have been contacted through the Brain Injury Association. Volunteers in this study need to have a family member who sustained an acquired brain injury (e.g., traumatic brain injury, stroke) after the age of 16. This family member must also have completed their acute rehabilitation following their brain injury. Additionally, volunteers must be comfortable using the English language and have no significant visual impairment.

Taking part in this study is voluntary.

It is your choice whether you take part or not. You do not have to take part in this study. Also, if you do agree to take part, you can stop taking part at any stage without having to give a reason. Whatever you decide to do, the health care your family member receives will not be affected. Please feel free to talk with a friend or family member if you need help to make the decision about whether to take part or not. Also, you are welcome to have a support person to be with you when we meet.

What will happen to the information collected in the study?

The things you say and the information we gather about your family member will be kept confidential and used for research purposes only. No material that identifies you, or your family member, will be used in any report on this project. We will store your family member's information in a secure location, and only those involved in the project will be able to see it. We will store your family member's information for seven years after the end of the project, after which time our records will be destroyed.

Finding out about the results of the study.

If you would like to find out the results of the project, please tick the first YES box on the consent form. After the project is completed, we will mail you the results.

Further information.

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington.

If you have any queries or concerns regarding your rights as a participant in this study you may wish to contact a Health and Disability Advocate, telephone:

- Auckland region 0800 555 050
- Wellington region 0800 42 36 38 (0800 4 ADNET)

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Mark Lewis	Duncan Babbage, PhD	Janet Leathem, PhD
Doctoral Candidate	Clinic Director	Professor of Neuropsychology

Appendix E

Consent Forms

Cognitive Disability after Brain Injury

Survey Consent Form

English	I wish to have an interpreter	Yes	No
Maori	E hiahia ana ahau ki tetahi kaiwhakamaori/kaiwhaka pakeha korero.	Ae	Kao
Cook Island	Ka inangaro au I tetai tangata uri reo.	Ae	Kare
Fijian	Au gadreva me dua e vakadewa vosa vei au.	Io	Sega
Niuean	Fia manako au ke fakaaoga e taha tagata fakahokohoko kupu.	E	Nakai
Samoan	Out e mana'o ia I ai se fa'amatala upu.	Ioe	Leai
Tokelaun	Ko au e fofou ki he tino ke fakaliliu te gagana Peletania ki no gagana o na motu o te Pahefika.	Ioe	Leai
Tonga	Oku ou fiema'u ha fakatonulea.	Io	Ikai

I have read and I understand the information sheet for volunteers taking part in this study.

The nature and purpose of the study have been explained to me. I have had the opportunity to discuss this study and ask questions about it. I am satisfied with the answers I have been given. I have had the opportunity to use family/whanau support or a friend to help me ask questions and understand the study. I have had time to consider whether to take part.

I understand the following.

- Taking part in this study is voluntary (my choice). I may withdraw from the study at any time and this will in no way affect my health care.
- I will be in control of what I do and what happens to me. I can ask questions or have a break when I need one.
- My participation in this study is confidential. No material that could identify me will be used in any published reports on this study.
- I would like to receive a copy of the results (tick one): Yes ☐ No ☐

  (There may be a long delay between when you take part and when the results are known).
- I am willing to be contacted about a follow-up study (tick one): Yes ☐ No ☐

I hereby agree to take part in this study.

My full name is:__________________________

Date:__________________________

Signature:__________________________

Project explained by:__________________________

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Cognitive Disability after Brain Injury

Survey Consent Form - Family Member Contact

English	I wish to have an interpreter	Yes	No
Maori	E hiahia ana ahau ki tetahi kaiwhakamaori/kaiwhaka pakeha korero.	Ae	Kao
Cook Island	Ka inangaro au I tetai tangata uri reo.	Ae	Kare
Fijian	Au gadreva me dua e vakadewa vosa vei au.	Io	Sega
Niuean	Fia manako au ke fakaaoga e taha tagata fakahokohoko kupu.	E	Nakai
Samoan	Out e mana'o ia I ai se fa'amatala upu.	Ioe	Leai
Tokelaun	Ko au e fofou ki he tino ke fakaliliu te gagana Peletania ki no gagana o na motu o te Pahefika.	Ioe	Leai
Tonga	Oku ou fiema'u ha fakatonulea.	Io	Ikai

I have read and I understand the information sheet for volunteers taking part in this study.

The nature and purpose of the study have been explained to me. I have had the opportunity to discuss this study and ask questions about it. I am satisfied with the answers I have been given. I have had the opportunity to use family/whanau support or a friend to help me ask questions and understand the study. I have had time to consider whether to take part.

I understand the following.

- I am willing to allow the family member named below to be contacted for this study.

- Taking part in this study is voluntary (mine and my family member's choice). Either of us may withdraw from the study at any time and this will in no way affect my health care.
- My family member's participation in this study is confidential. No material that could identify me, or my family member, will be used in any published reports on this study.

I hereby agree to allow a family member to be contacted for this study.

My full name is:__________________________

Date:__________________________

Signature:__________________________

Project explained by:__________________________

Family member's name:__________________________

Family member's contact details:__________________________

____

____

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Cognitive Disability after Brain Injury

Survey Consent Form–Family Member

English	I wish to have an interpreter	Yes	No
Maori	E hiahia ana ahau ki tetahi kaiwhakamaori/kaiwhaka pakeha korero.	Ae	Kao
Cook Island	Ka inangaro au I tetai tangata uri reo.	Ae	Kare
Fijian	Au gadreva me dua e vakadewa vosa vei au.	Io	Sega
Niuean	Fia manako au ke fakaaoga e taha tagata fakahokohoko kupu.	E	Nakai
Samoan	Out e mana'o ia I ai se fa'amatala upu.	Ioe	Leai
Tokelaun	Ko au e fofou ki he tino ke fakaliliu te gagana Peletania ki no gagana o na motu o te Pahefika.	Ioe	Leai
Tonga	Oku ou fiema'u ha fakatonulea.	Io	Ikai

I have read and I understand the information sheet for volunteers taking part in this study.

The nature and purpose of the study have been explained to me. I have had the opportunity to discuss this study and ask questions about it. I am satisfied with the answers I have been given. I have had the opportunity to use family/whanau support or a friend to help me ask questions and understand the study. I have had time to consider whether to take part.

I understand the following.

- Taking part in this study is voluntary (my choice). I may withdraw from the study at any time and this will in no way affect the health care received by my family member.

- I will be in control of what I do and what happens to me. I can ask questions or have a break when I need one.
- My participation in this study is confidential and that no material that could identify me, or my family member, will be used in any published reports on this study.
- I would like to receive a copy of the results (tick one): Yes ☐ No ☐

  (There may be a long delay between when you take part and when the results are known).

I hereby agree to take part in this study.

My full name is:____________________________

Date:____________________________

Signature:____________________________

Project explained by:____________________________

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Appendix F

Brownie Recipe

Ingredients:

2/3 cup margarine
1/2 cup sugar
1/2 cup brown sugar
2 large eggs
1 teaspoon vanilla extract
1 cup flour
1/3 cup unsweetened cocoa
1/4 teaspoon baking powder
1/8 teaspoon salt

Directions:

Coat a 20-cm circular microwavable dish with cooking spray. Place margarine in a medium microwavable bowl. Microwave, uncovered, at 100% power (700 watts) for 20 to 30 seconds until softened. Stir in sugar and brown sugar until blended. Add eggs and vanilla and beat until well blended. Add flour, cocoa powder, baking powder, and salt and stir just until moistened. Spread batter evenly in dish. Place dish in microwave oven.

Microwave, uncovered, at 100% power for 5 1/2 to 7 minutes until a wooden pick inserted in centre and around edges comes out clean. If the microwave does not rotate independently, rotate the dish one-fourth turn every 2 minutes (after 5 1/2 minutes of total cooking, check every 30 seconds). Cover dish with a tea towel and cool on a wire rack.

Appendix G

Changes to the Task Evidence Analysis Administration Manual

Words in plain text (e.g., text) have been retained from initial scoring manual and the added section is depicted in bold (e.g., **bold**).

Setting the Scene

Place all ingredients and utensils on a table or bench in the room where the cooking will take place. Before the specific instructions of the task are given to participants, they will inevitably realise that the task involves cooking. The specific task should not be mentioned until the instructions are given. However, the examiner should casually interact with participant. For example, if a participant asks what they are going to be cooking, the examiner may respond by lifting up the cocoa powder and saying, "I'll let you know in a few moments, but it has chocolate in it". This early interaction may help ease any trepidation felt by participants.

Instructions

Say to participants "In this task, you will need to make some brownies and a hot drink. You can make your own hot drink or you can make some tea, which I have provided. All other ingredients and utensils have also been provided. **Start by making the brownies. Make the hot drink while the brownies are cooking.** Any sort of strategy you currently use to complete a task of this nature, you're allowed to use. Here is a recipe for making the brownies". Then hand the recipe to the participants. Conclude the instructions by saying "Work as quickly and efficiently as possible. You may begin now."

Planning & Organisation

The examiner should not provide a plan nor organise the steps necessary to complete the task. If individuals ask for help in planning or organising the task, the examiner should respond by saying, "I would like to see if you can do this by yourself" or "Just do your best". However, if participants ask for the examiner to confirm that what they are doing is correct (e.g., "so, I should use this cooking spray to coat this bowl?"), the examiner should answer the question (e.g., "the recipe says to coat the dish, which is actually over there"). Herein lies an essential element in determining when to aid participants. Examiners should only help participants when participants specifically ask for help. There is a subtle difference between a participant saying "it does not say how much margarine to use" and one that says "how much margarine should I use?". The former participant made a statement, with no indication that they necessarily want input from the examiner, whereas the latter participant asked a question to obtain information. Examiners should only answer question posed by participants.

Initiation

Examiners should prompt individuals to perform goal-related behaviours if they display initiation difficulties. This prompt should occur if a lapse time of 1 minute has occurred without a task-related behaviour being performed. If an examiner initiation prompt is necessary, the examiner should say "Now that you have completed [state the last completed aspect of the task], what should you do next to complete the [brownie or tea]?".

Directed & Sustained Attention

When individuals become distracted, it may be necessary for the examiner to guide them back to the task at hand. This prompt should occur following a continuous 1 minute period of non-task related activity. If an examiner attention prompt is necessary, the examiner should say, “What are you working on now?” or “What should you be working on now?”.

Appendix H

Changes to the Task Evidence Analysis Scoring Manual

Words in plain text (e.g., text) have been retained from initial scoring manual, deleted sections of the initial scoring manual are represented by a line through the words (e.g., ~~words~~), and added sections are depicted in bold (e.g., **bold**).

Ability

Planning & Organisation

Planning and organisation occurs along a continuum from systematic to non-systematic/haphazard. Systematic performance occurs when individuals have a clear sense of what steps are needed to complete the task and organise the steps in an appropriate order so as to complete the task (Lezak et al., 2004). **A step is considered correct if it is the next uncompleted step in the sequence or any previously uncompleted step. Skipping a step that has been previously skipped does not count as another error.**

There are four sections of the task (coloured grey on page 2 of the record form) in which the steps need not be in sequential order to be scored correct. However, these sections need to be performed in the sequential order in which they appear on page 2. Additional steps are only scored as planning and organisation errors if they detract from the successful completion of the task. For example, additional stirring or cooking would not be considered errors, whereas adding canola oil would be considered a planning and organisation error. If an individual indicates that they have finished the task, but have uncompleted steps remaining, record the remaining uncompleted steps as planning and organisation errors. This is in contrast to uncompleted steps as a result of running out of time, in which the uncompleted steps are not scored as planning and organisation errors. A minimum

of 21 steps are necessary to complete the task according to the instructions. Additional steps are not included in the Planning and Organisation Total Score.

When an individual's performance is systematic, **that is, they perform the steps in the correct order (i.e., no errors occur),** a score of *Full Function* should be given. A score of *Effective Function With Some Limitations* should be given when ~~there are some systematic aspects of the successfully performed task~~ **individuals perform 1 to 5 errors**. ~~For example, individuals may have a clear sense of what steps are needed to complete a portion of the task, have a general idea of what steps are needed to complete the task, or organise the steps in such a way that, although the task is completed successfully, could been organised in a more appropriate way.~~ A score of *Achieves Some Functional Outcomes* is given when ~~there are some systematic aspects of an unsuccessfully performed task~~ **individuals perform 6 to 10 errors**. ~~For example, individual's may have a clear sense of what steps are needed to complete a portion of the task, have a general idea of what steps are needed to complete the task, or organise the steps in such a way that results in an unsuccessfully completed task. This score is also given when time expires without the task being completed, but again, some aspects of the task are completed.~~ A score of *Significant Disability* is given when ~~performance is non-systematic/haphazard~~ **individuals perform 11 to 15 errors**. ~~In other words, there is no clear sense of what steps are needed to complete the task and the organisation of the steps are inappropriate to successfully complete the task. However, the individual may successfully complete some aspects of the task.~~ A score of *Apparent Total Disability* should be given when performance is non-systematic/haphazard ~~and no aspects of the task are completed successfully~~. **This is indicated when more than 15 errors (see Figure 1).**

Executive Memory

Executive memory relates to task outcome and the ability of individuals to complete the assigned task. **For example, if individuals add ⅓ cup of sugar rather than ½ cup, this would be scored as an executive memory error regardless of when they actually added the sugar to the mixture. A measurement error occurs when there is an observable difference between the correct amount specified in the recipe and the amount added by an individual of about 10% (e.g., adding slightly more or less than 1 cup of flour is considered correct as long as the amount of flour added is approximately 1 cup). Additional steps not included in the instructions that detract from the successful completion of the task are scored as executive memory errors. However, as with planning and organisation errors, if an additional step is trivial and does not detract from the successful completion of the task, the step is not considered an error. For example, adding canola oil, or adding a listed ingredient a second time, would be considered an Executive Memory error, whereas adding ⅓ cup of chocolate chips would not be considered an error. It is important to remember that it is the final amount/behaviour that is scored. For instance, if an individual places 1 cup of margarine in the bowl, but removes ⅓ cup before melting it, an executive memory error is not scored as this behaviour was eventually correct. However, an error detection and correction would have been scored. If a step is not completed, the uncompleted step is not scored as an executive memory error.** ~~If time expires before the task is completed, the examiner needs to determine whether or not the performed behaviours would have resulted in the successful completion of the task.~~

A score of *Full Function* is awarded if ~~the task is successfully completed~~ **all aspects of the task are completed according to the instructions (i.e., individuals do not perform any Executive Memory errors)**. If time expires, a full function score is awarded if ~~the task would likely to have been completed, had the individual been given more time~~ **, up to that point, all of the aspects of the task have completed according to the instructions given**. A score of *Effective Function With Some Limitations* is given if ~~the outcome of the task is slightly different from what was asked~~ **between 1 to 5 errors occur**. **If** ~~In the case where~~ time expires, ~~this~~ **a *Effective Function With Some Limitations*** score is awarded if ~~the performed behaviours are slightly different to ones that would have been used to successfully complete the task~~**, up to that point, between 1 to 5 errors have been performed**. A score of *Achieves Some Functional Outcomes* is given if ~~the outcome of the task is moderately related to the task~~ **between 6 to 10 errors occur**. If time expires, ~~this~~ **a *Achieves Some Functional Outcomes*** score is awarded if ~~the performed behaviours are moderately related to the task~~**, up to that point, between 6 to 10 errors have been performed**. A score of *Significant Disability* is awarded if ~~the completed task is quite different to one assigned by the examiner~~ **11 to 15 errors occur**. If time expires ~~before the completion of the task, this~~**, a *Significant Disability*** score is awarded if ~~the performed behaviours are quite different to ones that would have been used to successfully complete the task~~, **up to that point, between 11 to 15 errors have been performed**. A score of *Apparent Total Disability* is awarded if ~~the completed task is in no way related to the assigned task~~ **more than 15 errors have occurred**. If time expires, ~~this~~ **a *Apparent Total Disability*** score is given if ~~all of the behaviours that are performed are in no way related to the assigned task~~**, up to that point, more than 15 errors have been performed (see Figure 1)**.

Initiation

Initiation can occur through self-initiation, examiner prompting, or a combination of the two. Care should be taken to distinguish initiation difficulties from planning and organisation difficulties; it may appear that individuals are unable to initiate a task, when in fact, they simply do not know what to do next to successfully complete the task. If an individual asks the examiner to prompt them, that is, they verbally express their plan and/or organisation and are simply wanting the examiner to help them initiate the next part of the plan, the prompt is considered a compensatory strategy and a score of full function should be given. The examiner's instruction to begin the task is not considered an examiner's prompt, unless the examiner instructs an individual to begin the task a second time. **If individuals respond verbally to an examiner prompt (e.g., "I should add sugar"), the examiner may ask them to initiate that step. Sometimes, the examiner prompt is sufficient to help individuals initiate the various aspects of the task. In both of these cases, only one initiation prompt is scored. A second initiation prompt is given, and scored, if a lapse time of 1 minute has passed following the initial initiation prompt without a task-related behaviour being performed. If a step is not completed, do not prompt the individual to initiate the uncompleted step after a lapse time of 1 minute has passed from the time the step should have been performed and do not score the uncompleted step.**

An individual will achieve the *Full Function* score if they are able to self-initiate each aspect of the task within an appropriate amount of time, taking into account any time needed to plan each step in completing the task. A score of *Effective Function With Some Limitations* is given when there ~~is minimal initiation prompting from the examiner; individuals are able to self-initiate most of the aspects of the task~~ **are 1 to 5**

initiation prompts from the examiner. A score of *Achieves Some Functional Outcomes* is given if the examiner prompts an individual to initiate ~~aspects of the task as much as the individual self-initiated aspects of the task; there is roughly a 50:50 ratio of self-initiation and examiner prompting~~ **6 to 10 aspects of the task**. A score of *Significant Disability* is given when ~~there is minimal self-initiation; the examiner provides most of the initiation prompts~~ **the examiner provides 11 to 15 initiation prompts**. A score of *Apparent Total Disability* is given when the examiner provides ~~all of the initiation prompts. That is, an individual never self-initiates aspects of the task~~ **more than 15 initiation prompts (see Figure 1)**.

Cognitive Shifting

Cognitive shifting refers to the ability to move from one aspect of the task to another aspect of the task. Scoring this criterion requires the examiner to distinguish between perseverative behaviours versus repeated behaviours that occur as a result of attentional deficits (Lezak et al., 2004**, for a fuller definition of cognitive shifting**). **Examples of difficulties in cognitive shifting may include continuing to mix the ingredients despite being well mixed and continuing to coat dish with cooking spray despite having already well coated it. If a step is not completed, the uncompleted step is not scored as either smooth or difficult.**

A *Full Function* score should be awarded if individuals are able to advance through the various aspects of the task without difficulty. That is, they stop performing a completed aspect of the task and begin performing the ensuing aspect of the task. A score of *Effective Function With Some Limitations* should be given if individuals ~~display perseverative behaviours following the completion of some of the~~ **have difficulty advancing through 1 to 5** aspects of the task. A score of *Achieves Some Functional*

Outcomes should be awarded if individuals ~~display many perseverative behaviours~~ **have difficulty advancing through 6 to 10 aspects of the task**. ~~In other words, following the completion of roughly half of the aspects of the task, individuals display perseverative behaviours~~. A score of *Significant Disability* should be given if individuals ~~display perseverative behaviours following the completion of most of the~~ **have difficulty advancing through 11 to 15** aspects of the task. A score of *Apparent Total Disability* should be awarded if individuals ~~display perseverative behaviours after the completion of each aspect of the task~~ **have difficulty advancing through more than 15 aspects of the task (see Figure 1)**.

~~Inhibition~~ **Impulsivity**

Scoring ~~inhibition~~ **impulsivity** requires the examiner to determine whether or not the individual is able to inhibit impulsive behaviours. **Examples of impulsive behaviours may include not reading the instructions before starting the task, adding additional ingredients (this must be distinguished from planning and organisation difficulties), not waiting for items to properly mix or cook, etc.** ~~This criteria becomes easier to score if a Likert scale, ranging from 0 to 4, is used regarding the presence of impulsive behaviours.~~

A score of *Full Function* should be given if ~~the~~ individual**s** do~~es~~ not display impulsive behaviours~~, thereby displaying the ability to inhibit impulsive responding (0)~~. A score of *Effective Function With Some Limitations* is given if ~~the~~ individual**s** display~~s~~ ~~some~~ **between 1 to 5** impulsive behaviours ~~(1)~~. A score of *Achieves Some Functional Outcomes* is given if ~~the~~ individual**s** display~~s many~~ **between 6 to 10** impulsive behaviours (2). A score of *Significant Disability* is given if ~~the majority of the individual's behaviours are impulsive (3)~~ **individuals display between 11 to 15**

impulsive behaviours. A score of *Apparent Total Disability* is given if ~~all of the behaviours that an individual performs are impulsive (4)~~ **more than 15 behaviours an individual performs are deemed to be impulsive (see Figure 1)**.

Directed & Sustained Attention

Directed and sustained attention refers to the ability to concentrate on the task, despite external or internal distractions. An examiner's prompt to attend to the task should not be scored as an examiner's initiation prompt. **An examiner's attention prompt is scored each time the examiner prompts a participant to re-attend to the task at hand. Individuals may be able to independently re-attend to the task following a period of less than 1 minute of non-task related activity. When this occurs, individuals have self prompted themselves back to the task and a self-attention prompt is scored.**

A score of *Full Function* should be given if individuals are able to attend to the task until it has been successfully completed, or until time expires **(i.e., there are no recorded examiner- or self- attention prompts)**. ~~This score should also be given when there are no examiner attention prompts.~~ A score of *Effective Function With Some Limitations* should be given ~~if, for the most part, individuals attend to the task at hand or there are few examiner attention prompts~~ **when 1 to 5 attention prompts occur either through self prompting, examiner prompting, or a combination of the two**. A score of *Achieves Some Functional Outcomes* should be given ~~if individuals have moderate difficulty concentrating of the task at hand~~ **when 6 to 10 attention prompts have occurred either through self prompting, examiner prompting, or a combination of the two**. ~~For example, an individual's concentration often shifts back and forth from task related stimuli to non-task related stimuli or after a period of sustained attention, an~~

~~individual is unable to re-attend to the task. Alternatively, this score should be given if the examiner provides a moderate amount of attention prompts.~~ A score of *Significant Disability* should be given ~~if individuals attended to non-task stimuli more than they attended to task-related stimuli or when the examiner provides many attention prompts~~ **when 11 to 15 attention prompts have occurred either through self prompting, examiner prompting, or a combination thereof**. A score of *Apparent Total Disability* should be given ~~if individuals are unable to sustain their attention to the task~~ **when more than 15 attention prompts have occurred either through self prompting, examiner prompting, or combination thereof (see Figure 1)**. ~~Alternatively, this score should be given when the examiner is constantly prompting individuals to attention to the task.~~

Error Detection

~~Scoring~~ **Correctly identifying** error detection can be extremely difficult. For example, the same behaviour may be displayed by individuals who do not detect errors and by individuals who do not correct a detected error because of pathological inertia (Lezak et al., 2004). Furthermore, examiners may believe that an undetected error has occurred, when in fact, the individual has simply organised the aspects of the task in an inappropriate manner. If participants indicate that they have detected an error (e.g., by saying “I shouldn’t have done that” or “That’s not right”), then scoring this criteria is made easier. However, **it is more likely that individuals will be more subtle. Thus, any discussion about the accuracy or inaccuracy of a particular decision counts as an error detection.** Examiners should not instruct participants to indicate when they detect errors as this might cue participants to monitor their behaviour for errors. **See error correction for additional error detection scoring criteria.**

There are four error detection scores. A False Alarm is scored when individuals believe they have detected an error, but have not performed one (e.g., an individual may say "There's no brown sugar here" when they cannot find it on the bench). A Monitoring score is given when individuals check to make sure their correct behaviour is indeed correct before continuing with the task (e.g., an individual may measure a ½ cup of sugar and then check the recipe to ensure that it calls for ½ cup of sugar). A Detection score is given when an individual performs an error and later realised they have made an error. A Failure score is given when an individual performs an error, but fails to detect the error.

The number of Detections and Failures made during the task determines which method is used to score Error Detection. When three or fewer Detections, Failures, or any combination of the two are made, the scoring criteria is as follows. A *Full Function* score is given when participants are able to detect all **(100%)** of the errors that occur. These errors may have occurred due to poor planning, organising, or implementation. When a participant completes the task without making an error, a *Full Function* score is given. ~~However, in this case, examiners should note that there was *No Evidence* for the score and that there is *No Confidence in Prediction*.~~ A**n** *Effective Function With Some Limitations* **score** should be awarded when ~~most of the errors are detected~~ **individuals are able to detect 50% or more of their errors, but fail to detect at least one error**. A score of *Achieves Some Functional Outcomes* is ~~achieved~~ **given** when ~~roughly half of a participant's errors are detected~~ **individuals detect between 25% and 49.99% of their errors**. A *Significant Disability* score is given when ~~most of the error go undetected~~ **individuals detect less than 25% of their errors**. An *Apparent Total Disability* score is ~~given when all of the errors that occur go~~

~~undetected~~ **not given when three or fewer errors have been detected as there is not enough information to make such a decision (see Figure 2)**.

When more than three Detections, Failures, or any combination of the two are made during the task, score this criteria is as follows. A *Full Function* score is given when participants are able to detect all (100%) of the errors that occur or when a participant completes the task without making an error. An *Effective Function With Some Limitations* score should be awarded when individuals are able to detect 75% or more of their errors, but fail to detect at least one error. A score of *Achieves Some Functional Outcomes* is given when individuals detect between 50% and 74.99% of their errors. A *Significant Disability* score is given when individuals detect between 25% to 49.99% of their errors. An *Apparent Total Disability* score is given when less than 25% of an individual's errors are detected (see Figure 2).

Error Correction

Correcting an error requires that an error first be detected. Thus, when an error correction occurs, an error detection should also be noted. An error correction occurs when, after an error has been detected, individuals change their behaviour. This may take the form of returning to a previously performed (correct) behaviour or modifying their behaviour to accommodate their current situation.

The number of errors detected during the task determines which method is used to score error correction. When three or fewer error are detected, the scoring criteria is as follows. A *Full Function* score is ~~awarded if individuals~~ **given when participants** are able to correct all **(100%)** of the errors that have been detected. [It should be noted that some individuals may complete the task, or run out of time, without having made an error. If this occurs, a *Full Function* score should be given] ~~as well as a *No Evidence* score and a *No Confidence in Prediction* score~~. A score of *Effective Function With Some Limitations* ~~is given~~ **should be awarded** when individuals correct ~~the majority of the errors that have been detected~~ **50% or more of their detected errors, but at least one detected error remains uncorrected**. A score of *Achieves Some Functional Outcomes* is ~~awarded~~ **given** when individuals correct ~~roughly 50 percent of the errors that had been detected~~ **between 25% and 49.99% of their detected errors**. A ~~score of~~ *Significant Disability* **score** is given when individuals ~~fail to~~ correct ~~the majority of the detected errors~~ **less than 25% of their detected errors**. An ~~score of~~ *Apparent Total Disability* **score** is ~~awarded~~ not given ~~no~~ **when three or fewer errors have been** detected ~~errors are corrected~~ **as there is not enough information to make such a decision (see Figure 2)**.

When more than three errors are detected during the task, score this criteria as follows. A *Full Function* score is given when participants are able to correct all (100%) of the errors that have been detected or when a participant completes the task without making an error. An *Effective Function With Some Limitations* score should be awarded when individuals are able to correct 75% or more of their detected errors, but at least one detected error remains uncorrected. A score of *Achieves Some Functional Outcomes* is given when individuals correct between 50% and 74.99% of their detected errors. A *Significant Disability* score is given when individuals correct between 25% to 49.99% of their detected errors. An *Apparent Total Disability* score is given when less than 25% of the errors an individual detects are corrected (see Figure 2).

~~Information Processing Speed~~ **Time-Management**

~~Information processing speed~~ **Time-management** refers to ~~time-management and~~ the ability to spend appropriate amounts of time on various aspects of the task (Boone et al., 1998; Lezak et al., 2004). ~~Each task can be completed in less than 20 minutes.~~ **The task can be completed in less than 30 minutes. However, individuals are given 40 minutes to complete the task before the task is abandoned.** ~~Therefore, tasks~~ **If the task is** completed ~~successfully~~ within ~~20~~ **30** minutes**,** ~~are scored as~~ **a score of** *Full Function* **is awarded**. A score of *Effective Function With Some Limitations* is awarded if individuals complete the task ~~successfully~~ within ~~30~~ **35** minutes. ~~If individuals fail to successfully complete the task within the allotted time, but successfully complete some aspects of the task,~~ A score of *Achieves Some Functional Outcomes* should be given **if individuals complete the task within 40 minutes**. A score of *Significant Disability* is awarded if individuals fail to complete the task within ~~the allotted time and had only completed roughly half of the~~ **40 minutes, but have completed more than 10** aspects of the task ~~that are needed to successfully complete the task~~ **(not including any additional steps)**. A score of *Apparent Total Disability* should be given if time expires **(i.e., more than 40 minutes)** and individuals have ~~only~~ completed ~~some of the~~ **10 or less** aspects of the task ~~that are needed for successful task completion~~ **(not including any additional steps; see Figure 1)**.

~~Evidence~~

~~The Evidence score is awarded based on the amount, or type, of evidence recorded. For each aspect of EF, the evidence is assessed to determine whether the Ability score is supported by participant behaviours. A score of *Unequivocal* should be awarded when the evidence clearly and indisputably supports the Ability score. For~~

~~example, an individual may advance through each aspect of the task smoothly (i.e., they never perseverated). In this case, an unequivocal score would be given as there is no doubt that the participant is able to make cognitive shifts. A *Strong Evidence* score should be given when most of the evidence, or there are several salient behaviours that, support the Ability score. This score may also be given when all of the evidence supports the Ability score, however, in the opinion of the examiner, it is of a lesser quality than unequivocal evidence. A *Moderate Evidence* score should be awarded when the majority of the evidence supports the Ability score. Alternatively, this score should be given if participants display behaviours that seem to supports the Ability score. A *Limited Evidence* score should be given when there is limited recorded evidence to support the Ability score. It may be that participants displayed one or two behaviours that suggests ability/disability. A *No Evidence* score should be awarded when there is no evidence to support the Ability score.~~

~~Confidence~~

~~A Confidence score is given based on the examiner's confidence that the evidence supports the Ability score. The scores range from *Very Strong Confidence* to *No Confidence in Prediction*.~~

Evidence of Capacity and Evidence of Incapacity

The Evidence of Capacity/Evidence of Incapacity score is a subjective assessment of a participant's performance. Evidence scores allow the examiner to couch the Ability score within their subjective impressions of a participant's overall performance. Evidence scores are based on a five-point Likert-type scale ranging from *Unequivocal* to *None*. When scoring Evidence of Capacity, a score of

Unequivocal should be given if the examiner has no doubt that a participant has the capacity to perform the specified aspect of executive functioning. A score of *Strong* should be awarded when the examiner believes strongly that a participant displays the capacity to perform the various aspects of executive functions. A score of *Moderate* is given when the examiner believes that a participant only displays a moderate capacity in perform the various aspects of executive function. A *Limited* score should be given when the examiner believes that a participant displays a limited capacity in performing the various aspects of executive functions. Finally, a *None* score should be awarded if the examiner believes that a participant displays no capacity in performing the various aspect of executive functioning.

When scoring Evidence of Incapacity, a score of *None* should be given if the examiner has no doubt that a participant has no incapacity in performing the specified aspect of executive functioning. A score of *Limited* should be awarded when the examiner believes that a participant displays limited incapacity in performing the various aspects of executive functioning. A score of *Moderate* is given when the examiner believes that a participants only displays moderate incapacity in performing the various aspects of executive function. A *Strong* score should be given when the examiner strongly believes that a participant displays the incapacity to performing the various aspects of executive functions. Finally, an *Unequivocal* score should be awarded when the examiner has no doubt that a participant has the incapacity to perform the various aspects of executive functioning.

Two examples might help clarify evidence scoring. Participant A performs all of the steps of the task in the correct order, thus achieving a Planning and Organisation score of *Full Function with/without Aids*. Based on Participant A's

performance, the examiner has no doubt that this participant has the capacity to plan and organise the steps necessary to complete the task and gives an Evidence of Capacity score of *Unequivocal*. The examiner also feels that Participant A did not display any signs of incapacity and gives an Evidence of incapacity score of *None*. This indicates that Participant A's ability to plan and organise is fully functioning, which is corroborated by the examiner's subjective assessment. Participant B performs three Executive Memory errors, thus achieving an Ability score of *Effective Function with Some Limitations*. However, these errors are such that they lead the examiner to believe that there is *Strong* evidence to suggest Incapacity and yet, there is *Strong* evidence to suggest Capacity (as 18 of the 21 steps were performed correctly). Thus, objectively (Ability score) Participant B is rated as *Effective Function with Some Limitations*, whereas subjectively (Evidence scores) Participant B is rated as being *Achieves Some Functional Outcomes*.

Scoring Page Two of the Record Sheet

For each aspect of the task, record Planning and Organisation (Plan), Executive Memory, Initiation, and Cognitive Shifting (Shifting). Under Plan, record the sequential order in which each aspect of the task is performed. There are 6 blank lines at the bottom of the page to record aspects of the task that are not included in the instructions. If an error occurs, mark the box to the right.

Under Executive Memory, record whether or not the aspect of the task is performed according to the instructions. If an error is recorded, note the error in the space provided. Do not record uncompleted steps as executive memory errors. Under Initiation, record whether each aspect is self-initiated or whether an examiner initiation prompt is necessary. Under Shifting, record whether the

transition from a previous aspect to the current aspect is performed smoothly or is performed with difficulty. The first aspect of the task is scored by monitoring the transition from not performing any aspects of the task (previous aspect) to performing the first aspect of the task (current aspect; see Table 1).

Impulsivity, Sustained and Directed Attention (Attention), Error Detection, and Error Correction should only be recorded when they occur (i.e., record the first incidence of impulsivity on the first line, the second incidence of impulsivity on the second line, etc.). Under Impulsivity, record a brief description or a one- or two- word description of the impulsive behaviour for future reference. When reviewing the tape at a later date, the actual behaviour can be recorded in the Impulsivity section on page 4 of the record sheet. Under Attention, record a brief or one- or two- word description when the individual becomes distracted for future reference. In addition, record whether they are able to self-initiate attention or whether an examiner attention prompt is given. Under Error Detection, record a short note to describe the False Alarm (FA), Monitoring (M), Detection (D), or Failure (F). Also, circle the appropriate category next to the note. Record error corrections only for those errors that are detected. Under Error Correction, record a short description of how the error was corrected for future reference and circle the 'C' to indicate that the error was corrected. If the detected error remains uncorrected, circle the 'F' under Error Correction (see Table 2).

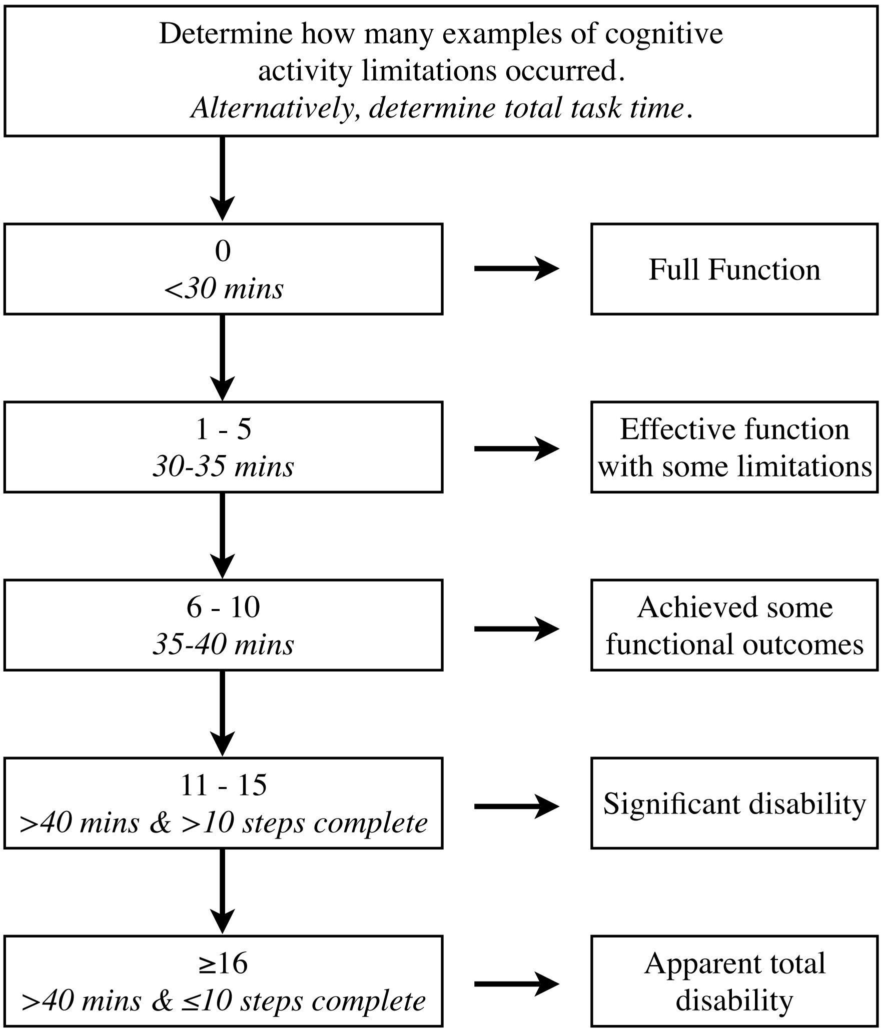

***Figure 1*. Decision tree for planning and organisation, executive memory, initiation, cognitive shifting, impulsivity, directed and sustained attention, and time management ability scores.**

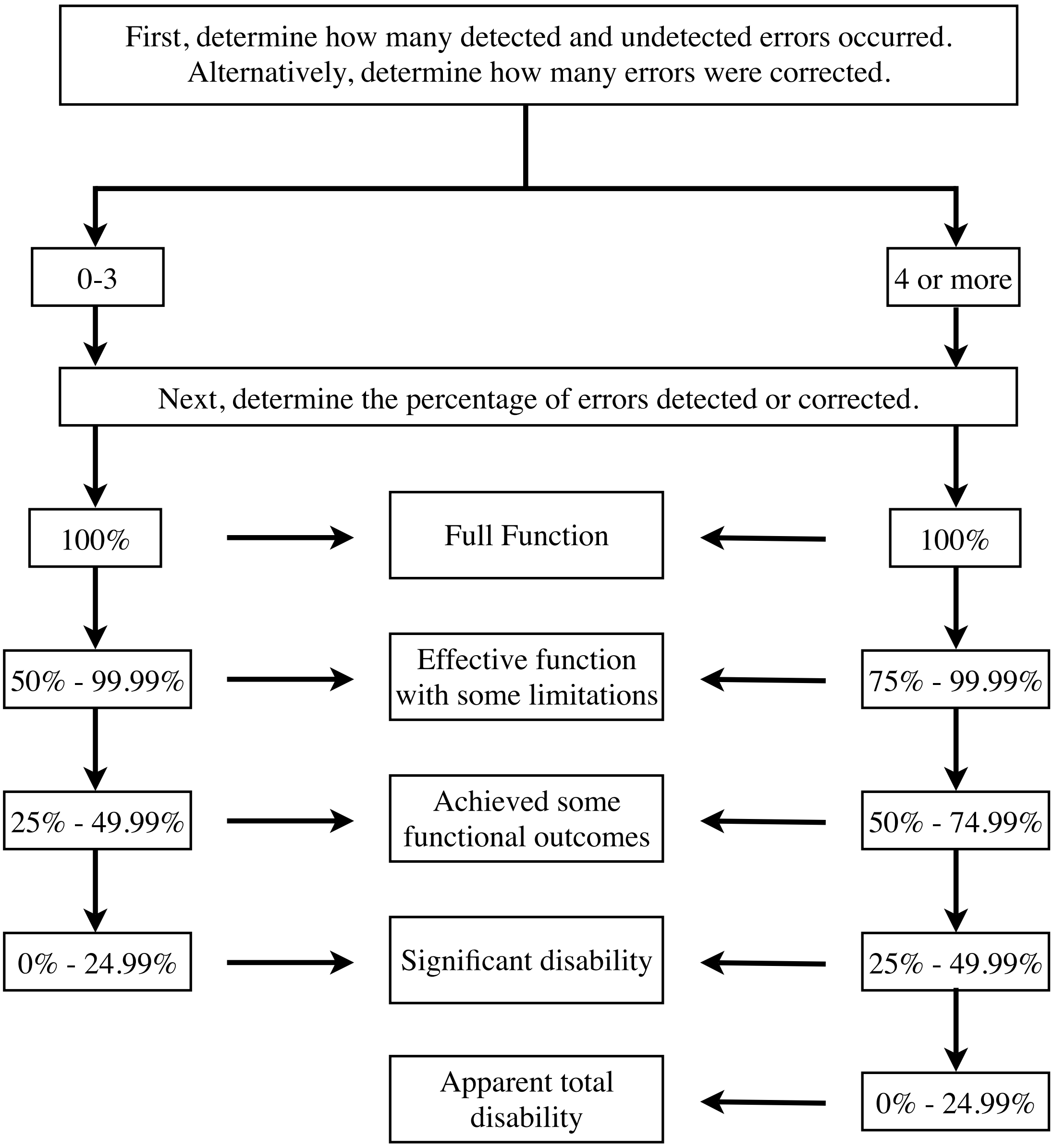

***Figure 2*. Decision tree for error detection and error correction ability scores.**

Table 1

Scoring Plan, Initiation, Shifting, and Executive Memory

Aspects of Task	Plan	Executive Memory (Correct/Error)	Initiation (Self/Examiner)	Shifting (Smooth/ Difficult)
1. Coat microwave dish	1 ☐	[✓ C] / E	[✓ Self] / Ex	[✓ S] / D
2. 2/3 c. margarine in bowl	5 ☐	C / [✓ E] 1/2 c.	Self / [✓ Ex]	[✓ S] / D
3. Microwave margarine	☑	C / E	Self / Ex	S / D
4. Add 1/2 c. sugar	4 ☐	[✓ C] / E	[✓ Self] / Ex	S / [✓ D]
5. Add 1/2 c. brown sugar	3 ☑	C / [✓ E] 3/4 c.	[✓ Self] / Ex	[✓ S] / D
6. Blend (beat or stir)	6 ☐	[✓ C] / E	Self / [✓ Ex]	S / [✓ D]
Add 1/2 c. canola oil	2 ☑	C / [✓ E] Addition	[✓ Self] / Ex	[✓ S] / D
Add 1/2 c. sugar	7 ☑	C / [✓ E] Addition	[✓ Self] / Ex	[✓ S] / D
Blend	8 ☐	[✓ C] / E	[✓ Self] / Ex	[✓ S] / D

Table 2

Scoring Impulsivity, Attention, Error Detection, and Error Correction

Impulsivity	Attention (Self/Examiner)	Error Detection (Detect/Fail)	Error Correction (Correct/Fail)
Cups	Sat down [✓ Self] / Ex	Margarine FA [✓ D] ➡ M F	Melt [✓ C] / F
Flour	Book Self / [✓ Ex]	Add flour FA [✓ D] ➡ M F	C / [✓ F]
Wash Dish	Cleaned Self / [✓ Ex]	"No salt" [✓ FA] D ➡ M F	C / F
	Self / Ex	Check tea FA D ➡ [✓ M] F	C / F
	Self / Ex	⅓ cup sugar FA D ➡ M [✓ F]	C / F

Appendix I

Scoring Manual Revisions Process

Ability Scores

Planning and Organisation

Planning and organisation errors were initially deemed to have occurred when individual steps of the task were completed out of order, as specified in the recipe (e.g., adding brown sugar before white sugar was considered an error as the recipe stated to add white sugar before adding brown sugar). After three participants had completed the functional task, it was deemed necessary to further clarify the order in which the steps were to be completed as there was some confusion regarding the scoring of steps that were skipped. For this reason, a step was considered correct if it was the next uncompleted step in the sequence or any previously uncompleted step, as this indicated that participants were adjusting their organisation of the task to accommodate an error. Skipping a step that was previously skipped did not count as another error.

Prior to the fourth participant completing the task, it was also concluded that within four specific sections of the task, the steps did not need to be completed in the specified order in order to be considered correct. However, these sections needed to be completed in the specified order. The reasoning behind this was that the outcome of the task was not affected if participants added brown sugar immediately before adding white sugar. Therefore, as the outcome of the task was not affected by their behaviour, their planning and organisation of the task should be considered to be correct. In addition to these modifications, only additional steps that detracted from the successful completion of the task were considered errors. This modification was added because, for example, stirring the mixture after adding white sugar, but before adding brown sugar, did not affect the outcome of the task and, therefore, should not be considered an error.

After the fourth participant had completed the functional task, it was deemed necessary to exclude additional steps in the calculation of the total planning and organisation score. Furthermore, errors were scored independently of the total score. Therefore, the maximum planning and organisation total score that could be achieved was 21, as there were 21 steps needed to complete the task. Additional steps that detracted from the successful completion of the task were considered errors and these errors, in addition to the errors from the 21 steps, contributed to the error score. After the final participant had completed the task, it was explicitly stated that omitted steps were considered errors.

Executive Memory

When a participant did not accurately complete a step of the task, they were deemed to have committed an executive memory error. This definition did not change throughout the manual revisions. Some leeway was given to all participants in that they did not need to measure each ingredient exactly. This leeway, which stated that the amount of an ingredient added should be within 10% of the specified amount, was explicitly stated after all participants had completed the task. The reasoning behind this was that many participants added approximately correct amounts of ingredients, but did not necessarily take the time to ensure that their measurements were exactly as stated in the recipe. After the second participant had completed the task, it was concluded that only additional tasks that detracted from the successful completion of the task should be scored as executive memory errors. The reason for this change was that the second participant stirred the mixture after adding the white sugar; the recipe states to stir the mixture after adding the brown sugar. As the outcome of the task was not affected by the second participant's additional stirring, the scoring criteria was adjusted to suit.

Initiation

Initiation scores did not change during the rating scale revision process as the scoring appeared relatively straightforward. Initiation scores were given according to whether or not an examiner initiation prompt was needed.

Cognitive Shifting

No changes were made to the initial scoring method for cognitive shifting. Each step of the task was rated based on a participant's ability to successfully initiate a subsequent step of the task. This definition was deemed to be relatively straightforward and easy to score.

Impulsivity

This aspect of the rating scale was initially termed 'inhibition'. However, as inhibition is an internal process and, therefore, unobservable, it was changed to impulsivity. No changes were made to actual scoring of impulsivity as it was thought that examiners were able to recognise impulsive behaviours were relatively easily.

Directed and Sustained Attention

Directed and sustained attention examiner prompts were given to all participants after a one minute period of non-task related behaviour. However, after the first participant had been scored, it was decided that participants could redirect themselves back to the task without needing an examiner's prompt. This was because this participant had been unavoidably interrupted several times during the task, but always returned to the task after the interruption abated. Therefore, a self prompt was scored from the second participant onward when individuals independently re-attended to the task following a period of less than 1 minute of non-task related activity.

Error Detection

There were initially two error detection scores—detection and failure. There was also just one method of scoring participant performance. However, after the first participant had completed the task it was concluded that two scoring methods were needed as it was thought that the score *Apparent Total Disability* would not be appropriate if participants made only a few errors. For this reason, error detections were rated based on the number of errors they had made.

The rating scale was modified again after the third participant had completed the task. It was noticed that participants often monitored their behaviour to ensure that they did not make an error. For example, the second participant read the recipe, then enter the cooking time into the microwave, and then check the recipe again to ensure that the time was correct before starting the microwave. Based on this behaviour, two more scoring criteria were created. False alarms were scored when participants claimed that a mistake had occurred, when in fact, no mistake had occurred and monitoring was scored when participants monitored their behaviour to ensure that they did not make a mistake. These latter two scoring criteria were not included in the error detection total score.

Error Correction

The error correction was similar to the error detection score in that there was initially only one scoring method regardless of the number of corrections made.

However, after the first participant had completed the task, it was decided that participants needed to have the opportunity to correct a number of detected errors before a score of *Apparent Total Disability* could be made. Therefore, from the second participant onward, two scoring criteria were created based on the number of error corrections that were possible during the task.

Time Management

Time management was initially termed 'information processing speed', but was change to 'time management' after the first participant had completed that task. The name change occurred to clarify what was actually being measured, as this aspect of EF was simply a measure of the speed in which the task was completed.

Functional Ratings

Functional ratings (e.g., *Full Function*) were awarded based on the number of errors that occurred. However, no specific number of errors were initially ascribed to a particular rating. Rather, it was left to the examiner to determine the functional rating based on a participant's performance. After the first participant had completed the task, it became clear that guidance was necessary in assigning functional ratings. Therefore, from the second participant onward, functional ratings were awarded based on the number of cognitive activity limitations performed. Functional ratings for planning and organisation, executive memory, initiation, cognitive shifting, impulsivity, and sustained and directed attention were scored based on the same criteria. More specifically, a score of *Full Function* was awarded if participants did not perform any cognitive activity limitations. A score of *Effective Function With Some Limitations* was given if participants performed one to five cognitive activity limitations. A score of *Achieves*

Some Functional Outcomes was given if participants performed 6 to 10 cognitive activity limitations. A score of *Significant Disability* was given if participants performed 11 to 15 cognitive activity limitations. Finally, a score of *Apparent Total Disability* was awarded if participants performed more than 15 cognitive activity limitations.

Functional ratings for error detection and error correction were also modified after the first participant had completed the task and were dependent on the number of errors that had been detected and corrected. When three or fewer error were made or corrected, the scoring criteria was as follows. A *Full Function* score was given when participants were able to detect or correct all (100%) of their errors. A score of *Effective Function With Some Limitations* was awarded when participants detected or corrected 50% or more of their errors, but at least one error remained undetected or uncorrected. A score of *Achieves Some Functional Outcomes* was given when participants detected or corrected between 25% and 49.99% of their errors. A *Significant Disability* score was given when participants detected or correct less than 25% of their errors. An *Apparent Total Disability* score was not given when three or fewer errors have been detected as it was concluded that there was not enough information to make such a decision.

When more than three errors were detected or corrected, the functional ratings were as follows. A *Full Function* score was given when participants were able to detect or correct all (100%) of their errors. An *Effective Function With Some Limitations* score was awarded when participants detected or corrected 75% or more of their errors, but at least one error remained undetected or uncorrected. A score of *Achieves Some Functional Outcomes* was given when between 50% and 74.99% of a participant's errors were detected or corrected. A *Significant Disability* score was given when participants detected or corrected between 25% to 49.99% of their errors. An *Apparent*

Total Disability score was given when less than 25% of the errors a participant made were detected or corrected.

The functional rating for time management were modified for the first three participants, with the functional rating remaining the same from the fourth participants onward. The modifications were as follows. The first participant needed to complete the task within 20 minutes in order to achieve functional rating of *Full Function* score. If they completed task within 30 minutes, they were given a score of *Effective Function With Some Limitations*. The other scores were given based on the number of steps completed if the participant exceeded the 30 minutes time limit. These time limits were chosen because one of the examiners was able to complete the task in less than 20 minutes.

The times used to score participant performance was modified for the second participant. They were allotted 25 minutes to complete the task for a *Full Function* score and were to be asked to stop the task if they exceeded 35 minutes. The second participant performed the task in less than 25 minutes, but the third participant took slightly longer than 35 minutes to complete the task. Therefore, from the fourth participant onward, the time management scores were changed. Participants were given 30 minutes to complete the task in order to achieve a *Full Function* score. This was in keeping with the suggested cooking time as specified in the recipe. Participants were given 35 minutes to complete the task in order to achieve a score of *Effective Function With Some Limitations* and were asked to stop the task after 40 minutes. The remaining two scores were given based on the number of steps that have been completed after being asked to stop the task.

Evidence Scores

Evidence scores were initially awarded based on the amount, or type, of evidence recorded. For each aspect of EF, the evidence was assessed to determine the extent to with the ability score was supported by participant behaviours. A score of *Unequivocal* was awarded when the evidence clearly and indisputably supports the ability score. For example, if a participant advanced through each aspect of the task smoothly (i.e., they never perseverated), an unequivocal score would be given as there was no doubt that the participant was able to make cognitive shifts. A *Strong Evidence* score was given when most of the evidence supported the ability score. This score was also be given when all of the evidence supported the ability score, however, in the opinion of the examiner, it was of a lesser quality than unequivocal evidence. A *Moderate Evidence* score was awarded when some of the evidence supported the ability score. Alternatively, this score was given if participants displayed behaviours that seemed to support the ability score. A *Limited Evidence* score was given when there was limited recorded evidence to support the ability score, such as if a participant displayed only one or two behaviours that suggested ability/disability. A *No Evidence* score was awarded when there was no evidence to support the ability score.

Evidence scores were calculated for the first four participants. However, this rating was abandoned due to confusion in the scoring process. At the root of this confusion lay the functional rating. On occasion, the raters believed that the functional rating awarded to participants represented the lowest end of their ability, whereas on other occasions the examiners believed that the functional rating awarded to participants represented their highest possible ability. For example, after completing the task Participant A was awarded with a functional rating of *Effective Function With Some Limitations*. Based on the recorded evidence, Examiner A felt that there was strong evidence to suggest that Participant A's performance was at least *Effective Function With Some Limitations*, if not *Full Function*. Examiner B also recorded *Effective Function With Some Limitations* as a functional rating as well as recording strong evidence, but felt that the evidence was such that the functional rating was the highest Participant A could perform (i.e., the functional rating was *Effective Function With Some Limitations*, if not *Achieves Some Functional Outcomes*). A solution could not be found to record this subtle difference in opinion and, thus, the evidence score was discontinued.

Confidence Scores

Confidence scores were given based on the examiner's subjective confidence that the ability score was correct. This score may or may not have be tied to, nor influenced by, the quality of objective evidence. The scores ranged from *Very Strong Confidence* to *No Confidence in Prediction*. Like the evidence score, confidence scores were only calculated for the first four participants. The reason for this was that after it became clear that the evidence score needed to be modified, providing confidence scores were put on hold until a full revision of the examiner's subjective assessments was completed.

Appendix J

Brownie Ratings

Participant Rating - Please rate your:					
1. Baking Experience:	1 (None)	2 (A Little)	3 (Moderate)	4 (A Lot)	5 (Expert)
2. Confidence in baking:	1 (No Confidence)	2 (A Little)	3 (Moderately)	4 (A Lot)	5 (Very Confident)
3. How well do you think you did:	1 (Poor)	2	3 (Moderate)	4	5 (Excellent)
4. With regard to the taste of the brownie, how satisfied are you with the outcome:	1 (Not Satisfied)	2 (A Little)	3 (Moderately)	4 (Quite)	5 (Very Satisfied)

Examiner Rating - Rate the taste of the brownie:	1 (I wish I hadn't had the first piece)	2 (I would not have another piece)	3 (It was OK)	4 (I was good)	5 (I couldn't have made better myself)

Appendix K

Performance Notes

Participant: s30 Examiner: Mark Date Tested: 7/11/2007 Venue: Home

1. Cooking time (margarine) = 20 sec
2. Cooking time (brownie) = 6 mins
3. Total task time = 18 min 33 sec
4. Was brownie cooked? Yes
5. Measurements: filled the ½ cup to about ⅔s full of margarine and then put about 1 tbsp into the ⅓ cup measure; filled about ⅔s of the ½ cup with sugar; filled about ⅔s of the ½ cup with brown sugar; used a teaspoon (but not the measuring teaspoon) to measure the vanilla, but the measurement was approximately correct; all other measurements correct.

6. She turned the kettle on after seeing me put the tea on the countertop.

Appendix L

Cognitive Disability after Brain Injury

Clinical Trial Information Sheet

You are invited to take part in a study to examine a new way of measuring cognitive disability. If you accept this invitation, you will be asked to perform several activities you might normally perform on a daily basis.

What does taking part involve?

If you chose to take part, we will meet for a total of about two hours. This could be one longer session or two or three shorter sessions. These meetings can take place either at the Psychology Clinic at Massey University in Wellington or at a place more convenient to you, but that is also suitable for the purposes of the project. This might include being able to complete the research at your home. You will be paid $20 in partial recognition of your costs, time, and effort in participating in this research.

What will people in the study be asked to do?

The purpose of this research is to develop a new method of assessing the ongoing problems people experience after brain injury. You will be asked to perform a series of everyday tasks. These might involve preparing food, organising activities, or finding information from different places.

You will be asked a series of questions. The questions are about the general activities you perform in an average day (e.g., cooking breakfast) and what kind of rehabilitation strategies (e.g., breaking tasks up into manageable parts) you use to help complete your activities. The questions also cover topics related to your social, physical, and emotional functioning in the past month.

You will also be asked to complete several more formal tasks that are routinely used with people with brain injury. These tasks are designed to measure your ability to plan and organise information, monitor your own behaviour, and remember things. You may even have completed some of these tasks in the past.

What information will researchers collect?

We will collect information by monitoring your ability to perform a series of everyday tasks and more formal tasks. We will videotape parts of our visit to enable us to cross-check our new assessment approach. You can ask to have the video recorder turned off at any time. The recordings will be deleted at the end of the study.

If you have recently completed any of our formal assessment tasks as part of a clinical assessment, we would prefer to use those results rather than administer the tests again. That would save you time, and provide better data for us. This would only happen with your consent. That specific data (not any other information) would then be provided to us by your clinician.

Are there any benefits or risks?

The study does not have any direct benefits for you and is not designed to improve your rehabilitation. The procedures used in the study have no known expected risks or harmful effects.

What sorts of people are taking part in this study?

You may have been invited to participate in the study because you have been a client of the Psychology Clinic at Massey University or Cavit ABI Rehabilitation. If so, a clinician indicated you could be a suitable person to participate in this study. Alternatively, you may have been contacted through the Brain Injury Association.

Volunteers in this study need to have had an acquired brain injury (e.g., traumatic brain injury, stroke) after the age of 16, have completed their acute rehabilitation following their brain injury, be comfortable using the English language, and have no significant visual impairment.

Taking part in this study is voluntary.

It is your choice whether you take part or not. You do not have to take part in this study. Also, if you do agree to take part, you can stop taking part at any stage without having to give a reason. Whatever you decide to do, the health care you receive will not be affected.

Please feel free to talk with a friend or family member if you need help to make the decision about whether to take part or not. Also, you are welcome to have a support person to be with you when we meet.

What will happen to the information collected in the study?

The things you say and the information we gather about you, including any video recordings, will be kept confidential and used for research purposes only. No material that identifies you will be used in any report on this project. We will store your information in a secure location, and only those involved in the project will be able to see it. We will store your information for seven years after the end of the project, after which time our records will be destroyed.

Finding out about the results of the study.

If you would like to find out the results of the project, please tick the YES box on the consent form. After the project is completed, we will mail you the results.

Further information.

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington.

If you have any queries or concerns regarding your rights as a participant in this study you may wish to contact a Health and Disability Advocate, telephone:

- Auckland region 0800 555 050
- Wellington region 0800 42 36 38 (0800 4 ADNET)

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Mark Lewis	Duncan Babbage, PhD	Janet Leathem, PhD
PhD Student	Clinic Director	Professor of Neuropsychology

Cognitive Disability after Brain Injury

Family Member

You are invited to take part in a project to find out what kinds of activities people with brain injury perform on a daily basis and what kind of rehabilitation strategies they use to help them complete those activities.

What does taking part involve?

If you chose to take part, we will meet for a total of about thirty minutes. This meeting can take place either at the Psychology Clinic at Massey University in Wellington or at a place more convenient to you, but that is also suitable for the purposes of the project. This might include being able to complete the research at your home.

What will people in the study be asked to do?

The purpose of this research is find out what kinds of activities people with brain injury perform on a daily basis and what kinds of rehabilitation strategies they use to help them complete those activities. You will be asked a series of questions about the general activities your family member performs in an average day (e.g., cooking breakfast) and what kind of rehabilitation strategies (e.g., breaking tasks up into manageable parts) are used to help complete those activities. The questions also cover topics related to the social, physical, and emotional functioning of your family member in the past month.

What information will researchers collect?

We will collect information through the use of two questionnaires.

Are there any benefits or risks?

The study does not have any direct benefits for your family member and is not designed to improve their rehabilitation. The procedures used in the study have no known expected risks or harmful effects.

What sorts of people are taking part in this study?

You may have been invited to participate in the study because your family member has been a client of the Psychology Clinic at Massey University or Cavit ABI Rehabilitation. If so, a clinician indicated you could be a suitable person to participate in this study. Alternatively, you may have been contacted through the Brain Injury Association. Volunteers in this study need to have a family member who sustained an acquired brain injury (e.g., traumatic brain injury, stroke) after the age of 16. This family member must also have completed their acute rehabilitation following their brain injury. Additionally, volunteers must be comfortable using the English language and have no significant visual impairment.

Taking part in this study is voluntary.

It is your choice whether you take part or not. You do not have to take part in this study. Also, if you do agree to take part, you can stop taking part at any stage without having to give a reason. Whatever you decide to do, the health care your family member receives will not be affected. Please feel free to talk with a friend or family member if you need help to make the decision about whether to take part or not. Also, you are welcome to have a support person to be with you when we meet.

What will happen to the information collected in the study?

The things you say and the information we gather about your family member will be kept confidential and used for research purposes only. No material that identifies you, or your family member, will be used in any report on this project. We will store your family member's information in a secure location, and only those involved in the project will be able to see it. We will store your family member's information for seven years after the end of the project, after which time our records will be destroyed.

Finding out about the results of the study.

If you would like to find out the results of the project, please tick the first YES box on the consent form. After the project is completed, we will mail you the results.

Further information.

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756,
Wellington.

If you have any queries or concerns regarding your rights as a participant in this study you may wish to contact a Health and Disability Advocate, telephone:

- Auckland region 0800 555 050
- Wellington region 0800 42 36 38 (0800 4 ADNET)

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Mark Lewis	Duncan Babbage, PhD	Janet Leathem, PhD
PhD Student	Clinic Director	Professor of Neuropsychology

Cognitive Disability after Brain Injury

Clinical Trial Consent Form

English	I wish to have an interpreter	Yes	No
Maori	E hiahia ana ahau ki tetahi kaiwhakamaori/kaiwhaka pakeha korero.	Ae	Kao
Cook Island	Ka inangaro au I tetai tangata uri reo.	Ae	Kare
Fijian	Au gadreva me dua e vakadewa vosa vei au.	Io	Sega
Niuean	Fia manako au ke fakaaoga e taha tagata fakahokohoko kupu.	E	Nakai
Samoan	Out e mana’o ia I ai se fa’amatala upu.	Ioe	Leai
Tokelaun	Ko au e fofou ki he tino ke fakaliliu te gagana Peletania ki no gagana o na motu o te Pahefika.	Ioe	Leai
Tonga	Oku ou fiema’u ha fakatonulea.	Io	Ikai

I have read and I understand the information sheet for volunteers taking part in this study.

The nature and purpose of the study have been explained to me. I have had the opportunity to discuss this study and ask questions about it. I am satisfied with the answers I have been given. I have had the opportunity to use family/whanau support or a friend to help me ask questions and understand the study. I have had time to consider whether to take part.

I understand the following.

- Taking part in this study is voluntary (my choice). I may withdraw from the study at any time and this will in no way affect my health care.
- I will be in control of what I do and what happens to me. I can ask questions or have
  a break when I need one.
- My participation in this study is confidential. No material that could identify me will be used in any published reports on this study.

- I consent to parts of my visit being videotaped (tick one): Yes ☐ No ☐
- I would like to receive a copy of the results (tick one): Yes ☐ No ☐

  (There may be a long delay between when you take part and when the results are known).

Complete only if you have recently completed a neuropsychological assessment:

- I consent to my clinician forwarding to the researchers my past test results for the formal measures used in this study (D-KEFS Tower, D-KEFS Sorting, Rivermead BMT, and/or EBIQ). (Clinician name: ______________________________)

  Yes ☐ No ☐

I hereby agree to take part in this study. My full name is:____________________________

Date:____________________________

Signature:____________________________

Project explained by:____________________________

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Cognitive Disability after Brain Injury

Survey Consent Form–Family Member

English	I wish to have an interpreter	Yes	No
Maori	E hiahia ana ahau ki tetahi kaiwhakamaori/kaiwhaka pakeha korero.	Ae	Kao
Cook Island	Ka inangaro au I tetai tangata uri reo.	Ae	Kare
Fijian	Au gadreva me dua e vakadewa vosa vei au.	Io	Sega
Niuean	Fia manako au ke fakaaoga e taha tagata fakahokohoko kupu.	E	Nakai
Samoan	Out e mana'o ia I ai se fa'amatala upu.	Ioe	Leai
Tokelaun	Ko au e fofou ki he tino ke fakaliliu te gagana Peletania ki no gagana o na motu o te Pahefika.	Ioe	Leai
Tonga	Oku ou fiema'u ha fakatonulea.	Io	Ikai

I have read and I understand the information sheet for volunteers taking part in this study.

The nature and purpose of the study have been explained to me. I have had the opportunity to discuss this study and ask questions about it. I am satisfied with the answers I have been given. I have had the opportunity to use family/whanau support or a friend to help me ask questions and understand the study. I have had time to consider whether to take part.

I understand the following.

- Taking part in this study is voluntary (my choice). I may withdraw from the study at any time and this will in no way affect the health care received by my family member.
- I will be in control of what I do and what happens to me. I can ask questions or have a break when I need one.

- My participation in this study is confidential and that no material that could identify me, or my family member, will be used in any published reports on this study.
- I would like to receive a copy of the results (tick one): Yes ☐ No ☐

  (There may be a long delay between when you take part and when the results are known).

I hereby agree to take part in this study.

My full name is:________________________

Date:________________________

Signature:________________________

Project explained by:________________________

If you would like more information about the study, please fell free to contact one of us:

• Mark Lewis	telephone: 801 0492
• Dr. Duncan Babbage	telephone: 801 5799 Ext. 62039
• Professor Janet Leathem	telephone: 801 5799 Ext. 62035

Our postal address is: Psychology Clinic, Massey University, Private Box 756, Wellington

This study has received ethical approval from the Multi-region Ethics Committee which reviews National and Multi regional studies.

Appendix M

Intraclass Correlation Coefficient Equation for the Two-Way Random Effects Model

The ICC calculation was based on the formula

$$ICC(2,1) = \frac{\text{subject variability}}{\text{subject variability } + \text{ observer variability } + \text{ random error variability}} \qquad (2)$$

(Rankin & Stokes, 1998, p. 198), which results in the equation

$$ICC(2,1) = \frac{BMS - EMS}{BMS + (k-1)EMS + k(RMS - EMS)/n} \qquad (3)$$

(Shrout & Fleiss, 1979, p. 423).

Appendix N

Task Evidence Analysis Scoring Manual

Ability

Planning & Organisation

Planning and organisation occurs along a continuum from systematic to non-systematic/haphazard. Systematic performance occurs when individuals perform all of the necessary steps in order to complete the task and organise those steps in an appropriate order so as to complete the task (Lezak et al., 2004). A step is considered correct if it is the next uncompleted step in the sequence or any previously uncompleted step. Skipping a step that has been previously skipped does not count as another error. However, if the preparation of the tea occurs out of the specified order, record this as a planning and organisation error as usual. The preparation of the tea is then considered in isolation to the preparation of the brownie. When returning to score the planning and organisation of the brownie preparation, a step is considered correct if it is the next uncompleted step in the sequence or any previously uncompleted step and skipping a step that has been previously skipped does not count as another error. Omitted steps are considered errors.

There are four sections of the task (coloured grey on page 2 of the record form) in which the steps need not be in sequential order to be scored correct. However, these sections need to be performed in the sequential order in which they appear on page 2. Additional steps are only scored as planning and organisation errors if they detract from the successful completion of the task. For example, additional stirring or cooking would not be considered errors, whereas adding canola oil would be considered a planning and organisation error. If an individual indicates that they have finished the task, but have uncompleted steps remaining, record the remaining uncompleted steps as planning and organisation errors. This is in contrast to uncompleted steps as a result of running out of time, in which the uncompleted steps are not scored as planning and organisation errors. A minimum of 21 steps are necessary to complete the task according to the instructions. Additional steps are not included in the Planning and Organisation Total Score.

When an individual's performance is systematic, that is, they perform the steps in the correct order (i.e., no errors occur), a score of *Full Function* should be given. A score of *Effective Function With Some Limitations* should be given when individuals perform 1 to 5 errors. A score of *Achieves Some Functional Outcomes* is given when individuals perform 6 to 10 errors. A score of *Significant Disability* should be given when individuals perform 11 to 15 errors. A score of *Apparent Total Disability* should be given when performance is non-systematic/haphazard. This is indicated when more than 15 errors (see Figure 1).

Executive Memory

Executive memory relates to task outcome and the ability of individuals to complete each aspect of the task correctly. For example, if individuals add ⅓ cup of sugar rather than ½ cup, this would be scored as an executive memory error regardless of when they actually added the sugar to the mixture. A measurement error occurs when there is an observable difference between the correct amount specified in the recipe and the amount added by an individual of about 10% (e.g., adding slightly more or less than 1 cup of flour is considered correct as long as the amount of flour added is approximately 1 cup). Additional steps not included in the instructions that detract from the successful completion of the task are scored as executive memory errors. However, as with planning and organisation errors, if an additional step is trivial and does not detract from the successful completion of the task, the step is not considered an error.

For example, adding canola oil, or adding a listed ingredient a second time, would be considered an Executive Memory error, whereas adding ⅓ cup of chocolate chips would not be considered an error. It is important to remember that it is the final amount/behaviour that is scored. For instance, if an individual places 1 cup of margarine in the bowl, but removes ⅓ cup before melting it, an executive memory error is not scored as this behaviour was eventually correct. However, an error detection and correction would have been scored. If a step is not completed, the uncompleted step is not scored as an executive memory error.

A score of *Full Function* is awarded if all aspects of the task are completed according to the instructions (i.e., individuals do not perform any Executive Memory errors). If time expires, a full function score is awarded if, up to that point, all of the aspects of the task have completed according to the instructions given. A score of *Effective Function With Some Limitations* is given if between 1 to 5 errors occur. If time expires, a *Effective Function With Some Limitations* score is awarded if, up to that point, between 1 to 5 errors have been performed. A score of *Achieves Some Functional Outcomes* is given if between 6 to 10 errors occur. If time expires, a *Achieves Some Functional Outcomes* score is awarded if, up to that point, between 6 to 10 errors have been performed. A score of *Significant Disability* is awarded if 11 to 15 errors occur. If time expires, a *Significant Disability* score is awarded if, up to that point, between 11 to 15 errors have been performed. A score of *Apparent Total Disability* is awarded if more than 15 errors have occurred. If time expires, a *Apparent Total Disability* score is awarded if, up to that point, more than 15 errors have been performed (see Figure 1).

Initiation

Initiation can occur through self-initiation, examiner prompting, or a combination of the two. Care should be taken to distinguish initiation difficulties from planning and organisation difficulties; it may appear that individuals are unable to initiate a task, when in fact, they do not know what to do next to successfully complete the task. If an individual asks the examiner to prompt them, that is, they verbally express their plan and/or organisation and are simply wanting the examiner to help them initiate the next part of the plan, the prompt is considered a compensatory strategy and a score of full function should be given. The examiner's instruction to begin the task is not considered an examiner's prompt, unless the examiner instructs an individual to begin the task a second time. If individuals respond verbally to an examiner prompt (e.g., "I should add sugar"), the examiner may ask them to initiate that step. Sometimes, the examiner prompt is sufficient to help individuals initiate the various aspects of the task. In both of these cases, only one initiation prompt is scored. A second initiation prompt is given, and scored, if a lapse time of 1 minute has passed following the initial initiation prompt without a task-related behaviour being performed. If a step is skipped or omitted, do not prompt the individual to initiate the skipped or omitted step and do not score the uncompleted step.

An individual will achieve the *Full Function* score if they are able to self-initiate each aspect of the task within an appropriate amount of time, taking into account any time needed to plan each step in completing the task. A score of *Effective Function With Some Limitations* is given when there are 1 to 5 initiation prompts from the examiner. A score of *Achieves Some Functional Outcomes* is given if the examiner prompts an individual to initiate 6 to 10 aspects of the task. A score of *Significant Disability* is given when the examiner provides 11 to 15 initiation prompts. A score of *Apparent Total Disability* is given when the examiner provides more than 15 initiation prompts (see Figure 1).

Cognitive Shifting

Cognitive shifting refers to the ability to move from one aspect of the task to another aspect of the task. Scoring this criterion requires the examiner to distinguish between perseverative behaviours and repeated behaviours that occur as a result of attentional deficits (see Lezak et al., 2004, for a fuller definition of cognitive shifting). Examples of difficulties in cognitive shifting may include continuing to mix the ingredients despite being well mixed and continuing to coat dish with cooking spray despite having already well coated it. If a step is not completed, the uncompleted step is not scored as either smooth or difficult.

A *Full Function* score should be awarded if individuals are able to advance through the various aspects of the task without difficulty. That is, they stop performing a completed aspect of the task and begin performing the ensuing aspect of the task. A score of *Effective Function With Some Limitations* should be given if individuals have difficulty advancing through 1 to 5 aspects of the task. A score of *Achieves Some Functional Outcomes* should be awarded if individuals have difficulty advancing through 6 to 10 aspects of the task. A score of *Significant Disability* should be given if individuals have difficulty advancing through 11 to 15 aspects of the task. A score of *Apparent Total Disability* should be awarded if individuals have difficulty advancing through more than 15 aspects of the task (see Figure 1).

Impulsivity

Scoring impulsivity requires the examiner to determine whether or not an individual displays impulsive behaviours. Examples of impulsive behaviours may include not reading the instructions before starting the task, adding additional ingredients (this must be distinguished from planning and organisation difficulties), not waiting for items to properly mix or cook, etc.

A score of *Full Function* should be given if individuals do not display impulsive behaviours. A score of *Effective Function With Some Limitations* is given if individuals display between 1 to 5 impulsive behaviours. A score of *Achieves Some Functional Outcomes* is given if individuals display between 6 to 10 impulsive behaviours. A score of *Significant Disability* is given if individuals display between 11 to 15 impulsive behaviours. A score of *Apparent Total Disability* is given if more than 15 behaviours an individual performs are deemed to be impulsive (see Figure 1).

Directed & Sustained Attention

Directed and sustained attention refer to the ability to concentrate on the task, despite external or internal distractions. An examiner's prompt to attend to the task should not be scored as an examiner's initiation prompt. An examiner's attention prompt is scored each time the examiner prompts a participant to re-attend to the task at hand. Individuals may be able to independently re-attend to the task following a period of less than 1 minute of non-task related activity. When this occurs, individuals have self prompted themselves back to the task and a self-attention prompt is scored.

A score of *Full Function* should be given if individuals are able to attend to the task until it has been successfully completed, or until time expires (i.e., there are no recorded examiner- or self- attention prompts). A score of *Effective Function With Some Limitations* should be given when 1 to 5 attention prompts occur either through self prompting, examiner prompting, or a combination of the two. A score of *Achieves Some Functional Outcomes* should be given when 6 to 10 attention prompts have occurred either through self prompting, examiner prompting, or a combination of the two. A score of *Significant Disability* should be given when 11 to 15 attention prompts have occurred either through self prompting, examiner prompting, or a combination thereof. A score of *Apparent Total Disability* should be given when more than 15 attention prompts have occurred either through self prompting, examiner prompting, or combination thereof (see Figure 1).

Error Detection

Correctly identifying error detection problems can be extremely difficult. For example, the same behaviour may be displayed by individuals who do not detect errors and by individuals who do not correct a detected error because of pathological inertia (Lezak et al., 2004). Furthermore, examiners may believe that an undetected error has occurred, when in fact, the individual has simply organised the aspects of the task in an inappropriate manner. If participants indicate that they have detected an error (e.g., by saying “I shouldn’t have done that” or “That’s not right”), then scoring this criteria is made easier. However, it is more likely that individuals will be more subtle. Thus, any discussion about the accuracy or inaccuracy of a particular decision counts as an error detection. Examiners should not instruct participants to indicate when they detect errors as this might cue participants to monitor their behaviour for errors. See error correction for additional error detection scoring criteria.

There are four error detection scores. A False Alarm is scored when individuals believe they have detected an error, but have not performed one (e.g., an individual may say “There’s no brown sugar here” when they cannot find it on the bench). A Monitoring score is given when individuals check to make sure their correct behaviour is indeed correct before continuing with the task (e.g., an individual may measure a ½ cup of sugar and then check the recipe to ensure that it calls for ½ cup of sugar). A Detection score is given when an individual performs an error and later realised they have made an error. A Failure score is given when an individual performs an error, but fails to detect the error.

The number of Detections and Failures made during the task determines which method is used to score Error Detection. When three or fewer Detections, Failures, or any combination of the two are made, the scoring criteria is as follows. A *Full Function* score is given when participants are able to detect all (100%) of the errors that occur. These errors may have occurred due to poor planning, organising, or implementation. When a participant completes the task without making an error, a *Full Function* score is given. An *Effective Function With Some Limitations* score should be awarded when individuals are able to detect 50% or more of their errors, but fail to detect at least one error. A score of *Achieves Some Functional Outcomes* is given when individuals detect between 25% and 49.99% of their errors. A *Significant Disability* score is given when individuals detect less than 25% of their errors. An *Apparent Total Disability* score is not given when three or fewer errors have been detected as there is not enough information to make such a decision (see Figure 2).

When more than three Detections, Failures, or any combination of the two are made during the task, score this criteria is as follows. A *Full Function* score is given when participants are able to detect all (100%) of the errors that occur or when a participant completes the task without making an error. An *Effective Function With Some Limitations* score should be awarded when individuals are able to detect 75% or more of their errors, but fail to detect at least one error. A score of *Achieves Some Functional Outcomes* is given when individuals detect between 50% and 74.99% of their errors. A *Significant Disability* score is given when individuals detect between 25% to 49.99% of their errors. An *Apparent Total Disability* score is given when less than 25% of an individual's errors are detected (see Figure 2).

Error Correction

Correcting an error requires that an error first be detected. Thus, when an error correction occurs, an error detection should also be noted. An error correction occurs when, after an error has been detected, individuals change their behaviour. This may take the form of returning to a previously performed (correct) behaviour or modifying their behaviour to accommodate their current situation.

The number of errors detected during the task determines which method is used to score error correction. When three or fewer error are detected, the scoring criteria is as follows. A *Full Function* score is given when participants are able to correct all (100%) of the errors that have been detected. [It should be noted that some individuals may completed the task, or run out of time, without having made an error. If this occurs, a *Full Function* score should be given.] A score of *Effective Function With Some Limitations* should be awarded when individuals correct 50% or more of their detected errors, but at least one detected error remains uncorrected. A score of *Achieves Some Functional Outcomes* is given when individuals correct between 25% and 49.99% of their detected errors. A *Significant Disability* score is given when individuals correct less than 25% of their detected errors. An *Apparent Total Disability* score is not given when three or fewer errors have been detected as there is not enough information to make such a decision (see Figure 2).

When more than three errors are detected during the task, score this criteria is as follows. A *Full Function* score is given when participants are able to correct all (100%) of the errors that have been detected or when a participant completes the task without making an error. An *Effective Function With Some Limitations* score should be awarded when individuals are able to correct 75% or more of their detected errors, but at least one detected error remains uncorrected. A score of *Achieves Some Functional Outcomes* is given when individuals correct between 50% and 74.99% of their detected errors. A *Significant Disability* score is given when individuals correct between 25% to 49.99% of their detected errors. An *Apparent Total Disability* score is given when less than 25% of the errors an individual detects are corrected (see Figure 2).

Time-Management

Time-management refers to the ability to spend appropriate amounts of time on various aspects of the task (Boone et al., 1998; Lezak et al., 2004). The task can be completed in less than 30 minutes. However, individuals are given 40 minutes to complete the task before the task is abandoned. If the task is completed within 30 minutes, a score of *Full Function* is awarded. A score of *Effective Function With Some Limitations* is awarded if individuals complete the task within 35 minutes. A score of *Achieves Some Functional Outcomes* should be given if individuals complete the task within 40 minutes. A score of *Significant Disability* is awarded if individuals fail to complete the task within 40 minutes, but have completed more than 10 aspects of the task (not including any additional steps). A score of *Apparent Total Disability* should be given if time expires (i.e., more than 40 minutes) and individuals have completed 10 or less aspects of the task (not including any additional steps; see Figure 1).

Evidence of Capacity and Evidence of Incapacity

The Evidence of Capacity/Evidence of Incapacity score is a subjective assessment of a participant's performance. Evidence scores allow the examiner to couch the Ability score within their subjective impressions of a participant's overall performance. Evidence scores are based on a five-point Likert-type scale ranging from *Unequivocal* to *None*. When scoring Evidence of Capacity, a score of *Unequivocal* should be given if the examiner has no doubt that a participant has the capacity to perform the specified aspect of executive functioning. A score of *Strong* should be awarded when the examiner believes strongly that a participant displays the capacity to perform the various aspects of executive functions. A score of *Moderate* is given when the examiner believes that a participant only displays a moderate capacity in perform the various aspects of executive function. A *Limited* score should be given when the examiner believes that a participant displays a limited capacity in performing the various aspects of executive functions. Finally, a *None* score should be awarded if the examiner believes that a participant displays no capacity in performing the various aspect of executive functioning.

When scoring Evidence of Incapacity, a score of *None* should be given if the examiner has no doubt that a participant has no incapacity in performing the specified aspect of executive functioning. A score of *Limited* should be awarded when the examiner believes that a participant displays limited incapacity in performing the various aspects of executive functioning. A score of *Moderate* is given when the examiner believes that a participants only displays moderate incapacity in performing the various aspects of executive function. A *Strong* score should be given when the examiner strongly believes that a participant displays the incapacity to performing the various aspects of executive functions. Finally, an *Unequivocal* score should be awarded when the examiner has no doubt that a participant has the incapacity to perform the various aspects of executive functioning.

Two example might help clarify evidence scoring. Participant A performs all of the steps of the task in the correct order, thus achieving a Planning and Organisation score of *Full Function with/without Aids*. Based on Participant A's performance, the examiner has no doubt that this participant has the capacity to plan and organise the steps necessary to complete the task and gives an Evidence of Capacity score of *Unequivocal*. The examiner also feels that Participant A did not display any signs of incapacity and gives an Evidence of incapacity score of *None*. This indicates that Participant A's ability to plan and organise is fully functioning, which is corroborated by the examiner's subjective assessment. Participant B performs three Executive Memory errors, thus achieving an Ability score of *Effective Function with Some Limitations*. However, these errors are such that they lead the examiner to believe that there is *Strong* evidence to suggest Incapacity and yet, there is *Strong* evidence to suggest Capacity (as 18 of the 21 steps were performed correctly). Thus, objectively (Ability score) Participant B is rated as *Effective Function with Some Limitations*, whereas subjectively (Evidence scores) Participant B is rated as being *Achieves Some Functional Outcomes*.

Scoring Page Two of the Record Sheet

For each aspect of the task, record Planning and Organisation (Plan), Executive Memory, Initiation, and Cognitive Shifting (Shifting). Under Plan, record the sequential order in which each aspect of the task is performed. There are 6 blank lines at the bottom of the page to record aspects of the task that are not included in the instructions. If an error occurs, mark the box to the right.

Under Executive Memory, record whether or not the aspect of the task is performed according to the instructions. If an error is recorded, note the error in the space provided. Do not record uncompleted steps as executive memory errors. Under Initiation, record whether each aspect is self-initiated or whether an examiner initiation prompt is necessary. Under Shifting, record whether the transition from a previous aspect to the current aspect is performed smoothly or is performed with difficulty. The first aspect of the task is scored by monitoring the transition from not performing any aspects of the task (previous aspect) to performing the first aspect of the task (current aspect; see Table 1).

Impulsivity, Sustained and Directed Attention (Attention), Error Detection, and Error Correction should only be recorded when they occur (i.e., record the first incidence of impulsivity on the first line, the second incidence of impulsivity on the second line, etc.). Under Impulsivity, record a brief description or a one- or two- word description of the impulsive behaviour for future reference. When reviewing the tape at a latter date, the actual behaviour can be recorded in the Impulsivity section on page 4 of the record sheet. Under Attention, record a brief or one- or two- word description when the individual becomes distracted for future reference. In addition, record whether they are able to self-initiate attention or whether an examiner attention prompt is given. Under Error Detection, record a short note to describe the False Alarm (FA), Monitoring (M), Detection (D), or Failure (F). Also, circle the appropriate category next to the note. Record error corrections only for those errors that are detected. Under Error Correction, record a short description of how the error was corrected for future reference and circle the ‘C’ to indicate that the error was corrected. If the detected error remains uncorrected, circle the ‘F’ under Error Correction (see Table 2).

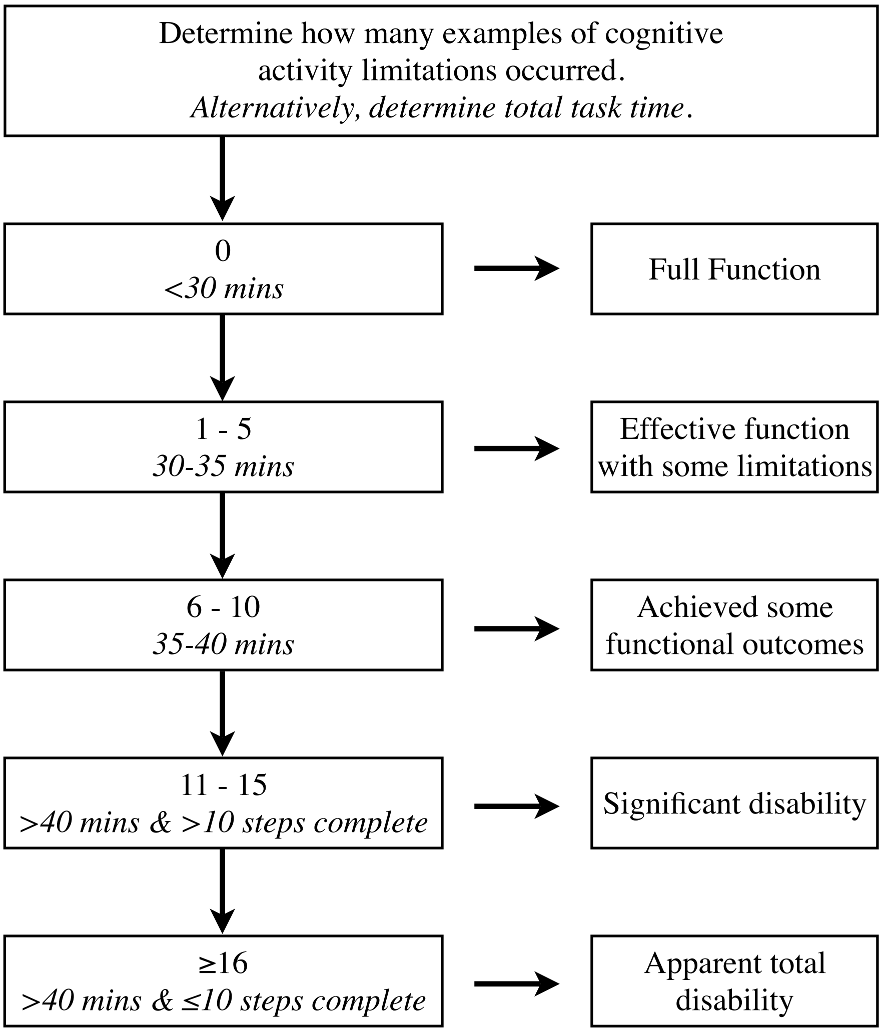

Figure 1. Decision tree for planning and organisation, executive memory, initiation, cognitive shifting, impulsivity, directed and sustained attention, and time management ability scores.

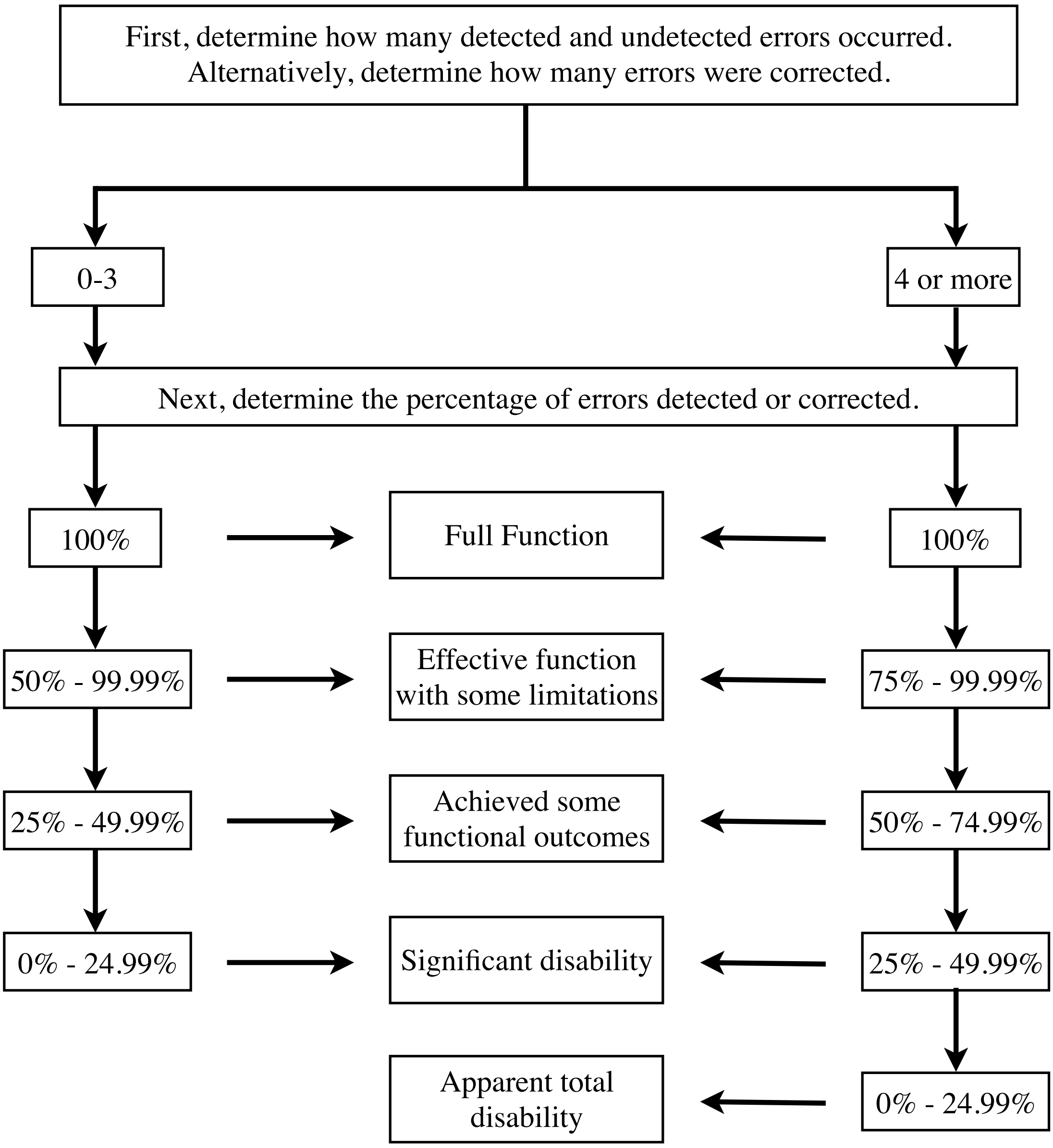

Figure 2. Decision tree for error detection and error correction ability scores.

Table 1

Scoring Planning and Organisation, Executive Memory, Initiation, and Cognitive Shifting

Aspects of Task	Plan		Executive Memory (Correct/Error)		Initiation (Self/Examiner)	Shifting (Smooth/ Difficult)
1. Coat microwave dish	1	☐	[✓ C] / E		[✓ Self] / Ex	[✓ S] / D
2. 2/3 c. margarine in bowl	5	☐	C / [✓ E]	1/2 c.	Self / [✓ Ex]	[✓ S] / D
3. Microwave margarine		☑	C / E		Self / Ex	S / D
4. Add 1/2 c. sugar	4	☐	[✓ C] / E		[✓ Self] / Ex	S / [✓ D]
5. Add 1/2 c. brown sugar	3	☑	C / [✓ E]	3/4 c.	[✓ Self] / Ex	[✓ S] / D
6. Blend (beat or stir)	6	☐	[✓ C] / E		Self / [✓ Ex]	S / [✓ D]
Add 1/2 c. canola oil	2	☑	C / [✓ E]	Addition	[✓ Self] / Ex	[✓ S] / D
Add 1/2 c. sugar	7	☑	C / [✓ E]	Addition	[✓ Self] / Ex	[✓ S] / D
Blend	8	☐	[✓ C] / E		[✓ Self] / Ex	[✓ S] / D

Table 2

Scoring Impulsivity, Sustained and Directed Attention, Error Detection, and Error Correction

Impulsivity	Attention (Self/Examiner)		Error Detection (Detect/Fail)		Error Correction (Correct/Fail)	
Cups	Sat down	[✓ Self] / Ex	Margarine	FA [✓ D] ➡ M F	Melt	[✓ C] / F
Flour	Book	Self / [✓ Ex]	Add flour	FA [✓ D] ➡ M F		C / [✓ F]
Wash Dish	Cleaned	Self / [✓ Ex]	"No salt"	[✓ FA] D ➡ M F		C / F
		Self / Ex	Check tea	FA D ➡ [✓ M] F		C / F
		Self / Ex	⅓ cup sugar	FA D ➡ M [✓ F]		C / F

Appendix O

1 - Task Instructions, Ability & Rating Scores **2** - Main Rating Page **3** - Planning & Organising, Executive Memory, Initiation, Cognitive Shifting, Impulsivity **4** - Attention, Error Detection & Correction, Time Management, Notes

Participant: ____________ Examiner: ____________ Date Tested: ____________ Venue: ____________

Task Instructions

In this task, you will need to make some brownies and a hot drink. You can make your own hot drink or you can make some tea, which I have provided. All other ingredients and utensils have also been provided. Start by making the brownies. Make the hot drink while the brownies are cooking. Any sort of strategy you currently use to complete a task of this nature, you're allowed to use. Here is a recipe for making the brownies [Hand it to participant]. Work as quickly and efficiently as possible. You may begin now.

If asked to help plan and organise the task: "I would like to see if you can do this by yourself" or "Just do your best".
Initiation Prompt: "Now that you have completed [state the last completed aspect of the task], what should you do next?"
Attention Prompt: "What are you working on now?" or "What should you be working on now?".
If asked whether or not an aspect has been completed: "I think you have, but it is up to you whether you add it or not". Alternatively, "I don't think you have, but it is up to you whether you add it or not".

Aspects of Task	Planning & Organisation (Total/Error)	Executive Memory (Error)	Initiation (Examiner)	Cognitive Shifting (Difficult)	Impulsivity	Attention (Self/Examiner)	Error Detection (FA/Monitor/Detect/Fail)	Error Correction (Correct/Fail)
Totals:	/					/	/ / /	/
Time Completed:							%	%

Examiner Rating - Rate the taste of the brownie:	1 (I wish I hadn't had the first piece)	2 (I would not have another piece)	3 (It was OK)	4 (It was good)	5 (I couldn't have made better myself)

Participant Rating - Please rate your:

1. Baking Experience:	1 (None)	2 (A Little)	3 (Moderate)	4 (A Lot)	5 (Expert)
2. Confidence in baking:	1 (No Confidence)	2 (A Little)	3 (Moderately)	4 (A Lot)	5 (Very Confident)
3. How well do you think you did:	1 (Poor)	2	3 (Moderate)	4	5 (Excellent)
4. With regard to the taste of the brownie, how satisfied are you with the outcome:	1 (Not Satisfied)	2 (A Little)	3 (Moderately)	4 (Quite)	5 (Very Satisfied)

1 - Task Instructions, Ability & Rating Scores **2** - Main Rating Page **3** - Planning & Organising, Executive Memory, Initiation, Cognitive Shifting, Impulsivity **4** - Attention, Error Detection & Correction, Time Management, Notes

Aspects of Task	Plan	Executive Memory (Correct/Error)	Initiation (Self/Examiner)	Shifting (Smooth/Difficult)	Impulsivity	Attention (Self/Examiner)	Error Detection (Detect/Fail)	Error Correction (Correct/Fail)
1. Coat microwave dish	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
2. 2/3 c. margarine in bowl	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
3. Microwave margarine (20-30 sec)	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
4. Add 1/2 c. sugar	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
5. Add 1/2 c. brown sugar	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
6. Blend (beat or stir)	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
7. Add 2 eggs	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
8. Add 1 tsp vanilla	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
9. Blend (beat or stir)	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
10. Add 1 c. flour	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
11. Add 1/3 c. cocoa	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
12. Add 1/4 tsp baking powder	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
13. Add 1/8 tsp salt	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
14. Stir until moistened	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
15. Spread evenly in dish	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
16. Cook (5 1/2 to 7 mins)	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
17. Boil water	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
18. Combine tea and water	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
19. Check with toothpick	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
20. Cover dish	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
21. Place on rack	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F
	☐	C / E	Self / Ex	S / D		Self / Ex	FA D M F ➡	C / F

Planning and Organisation	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Executive Memory	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Initiation	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Cognitive Shifting	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Implusivity	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Directed and Sustained Attention	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Error Detection	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Error Correction	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Time Management	**Evidence of capacity**	**=**	unequivocal	strong	moderate	limited	none
	Ability	**=**	Full function (with/without aids)	Effective function with some limitations	Achieves some functional outcomes	Significant disability	Apparent total disability
	Evidence of incapacity	**=**	none	limited	moderate	strong	unequivocal
Notes:							

Additional Notes: