Developing a Culture Fair Cognitive Estimation Test

& Clinical Research Portfolio

Volume 1
(Volume 2 bound separately)

Cathy Tran, BSc Honours

September 2015

Submitted in partial fulfillment of the requirements for the degree of Doctorate in Clinical Psychology (DClinPsy)

Institute of Health and Wellbeing
College of Medical, Veterinary and Life Sciences
University of Glasgow September 2015

University of Glasgow | Institute of Health & Wellbeing

Declaration of Originality Form

This form **must** be completed and signed and submitted with all assignments.

Please complete the information below (using BLOCK CAPITALS).

Name	Cathy Tran
Student Number	2019731t
Course Name	Doctorate in Clinical Psychology
Assignment Number/Name	Clinical Research Portfolio

An extract from the University's Statement on Plagiarism is provided overleaf. Please read carefully THEN read and sign the declaration below.

I confirm that this assignment is my own work and that I have:

Read and understood the guidance on plagiarism in the Doctorate in Clinical Psychology Programme Handbook, including the University of Glasgow Statement on Plagiarism	☑
Clearly referenced, in both the text and the bibliography or references, **all sources** used in the work	☑
Fully referenced (including page numbers) and used inverted commas for **all text quoted** from books, journals, web etc. (Please check the section on referencing in the 'Guide to Writing Essays & Reports' appendix of the Graduate School Research Training Programme handbook.)	☑
Provided the sources for all tables, figures, data etc. that are not my own work	☑
Not made use of the work of any other student(s) past or present without acknowledgement. This includes any of my own work, that has been previously, or concurrently, submitted for assessment, either at this or any other educational institution, including school (see overleaf at 31.2)	☑
Not sought or used the services of any professional agencies to produce this work	☑
In addition, I understand that any false claim in respect of this work will result in disciplinary action in accordance with University regulations	☑

DECLARATION:

I am aware of and understand the University's policy on plagiarism and I certify that this assignment is my own work, except where indicated by referencing, and that I have followed the good academic practices noted above

Signature CATHY TRAN Date 18.9.15

Acknowledgements

Firstly I would like to say a huge thank you to my research supervisors, Professor Jon Evans and Dr. Brian O'Neill for their constant encouragement, guidance and support throughout the duration of this research project. I greatly appreciate their patience and enthusiasm. They have made this process a great learning experience.

I must also say a big thank you to the staff at the Graham Anderson House who have been extremely helpful and have gone out of their way to help with the recruitment process as well as making me feel welcome at the unit. I thoroughly enjoyed meeting the team and service users. Similarly, thank you to Chloe Lack at Daniel Yorath House, who has also been very helpful and given her time to help with recruitment, and to Matthew Jamieson for his support with the review. Thanks also to all those who freely gave their time to take part in the research.

Thanks to my fellow trainees, for being so supportive over the past few years. I couldn't have done this without the support and fun they have brought to the training. Thank you to all my friends, my parents and my brother, who have been amazing support throughout, and have always listened and provided encouragement when I needed it most.

I would also like to thank NHS Education for Scotland and the University of Glasgow for providing the funding to complete my training.

Table of Contents

Volume I (this bound volume)

Chapter One: Systematic Review

The Validity of Cognitive Estimation Tests: A Systematic Review

Author: Cathy Tran [1]

[1] Mental Health and Wellbeing, University of Glasgow

Correspondence Address:
Mental Health and Wellbeing University of Glasgow
Gartnaval Royal Hospital
1055 Great Western Road Glasgow
G12 0XH
Email: c.tran.1@research.gla.ac.uk

Prepared in accordance with submission requirements for Journal of the International Neuropsychological Society *(See Appendix 1)*

Word Count: 5965 *(not including Abstract, Tables, Figures or References)*

Abstract

Background: Judgement, estimation and problem-solving abilities involve complex cognitive functions. Cognitive estimation tests (CETs) were developed to assess problem-solving abilities and were found to be a marker of frontal lobe function. Although CETs are used in everyday clinical practice, there is uncertainty as to whether they are useful, clinically valid tests.

Aims: The aim of this review was to collate evidence from studies assessing the validity of CETs, to offer an indication of both methodological quality and quality of reporting, and to suggest directions for future research in this area.

Methods: Correlational studies and case-control studies were included in this review. Studies were selected that examined associations between the CET and other executive functioning tests or compared a group of patients with a frontal or executive deficit, with a group of patients with brain injury not specific to the frontal lobes or those deemed not to have executive functioning deficits. Studies were rated according to the STROBE-22 reporting guidelines and an additional eight questions were used to assess methodological quality of the studies.

Main results: Twenty-one articles were included. Eight correlational studies, 11 case-control studies, and two of mixed design. Studies were rated as high, moderate and low quality for each checklist, then categorised into groups from highest (A) to lowest quality (E) based on both quality checklists. Regardless of overall quality, the majority of studies comparing patients with frontal deficits with patients with brain injury not specific to the frontal lobes or those deemed not to have executive functioning deficits found either no significant differences on CET performance (n=7) or found results indicating that the frontal group performed better than the comparison group (n=3). Correlational study results were varied, often showing poor construct validity of CETs.

Overall, results suggest that the CET is not a specific measure of executive functioning, and appears to reflect more general cognitive abilities. The CET also demonstrated little correlation with tests associated with everyday functioning, indicating poor ecological validity.

Conclusions: CETs do not appear to be effective measures of executive functions and do not predict everyday decision-making. It may be that the CET draws upon a number of cognitive functions, including those controlled by the frontal lobes, but not specific to this area. However, the limited number of high quality studies makes it difficult to draw firm conclusions, or generalise findings to other patient populations. There is a need for rigorously reported studies, as well as higher quality methodologies in studies, particularly with regards to statistical power, ensuring statistically sound methods to measure outcomes, and clearer rationales for selection of patient groups.

Introduction

Judgement, or the ability to make considered decisions or come to sensible conclusions, is important for effective independent living. Judgement is the result of a process by which evidence is evaluated, chances of different outcomes assessed, and an action decided (Blanchette & Richards, 2010). The process of estimation is considered to draw upon similar processes. To estimate requires the ability draw upon existing knowledge, use that information to generate possible answers (estimates), weigh up the possibilities and select the best option. Estimation therefore may be considered a relatively specific judgement task. If estimation abilities are compromised, the ability to make reasoned decisions may be affected. Impairment of estimation abilities as a result of brain injury can have serious implications for the individual and are important to assess. Assessment of an individual's ability to estimate may be useful in determining whether someone has the capacity to make important decisions. An injury may mean they are unable to draw upon relevant knowledge, understand consequences, or take time to think through a decision. Therefore, they may be more at risk of making decisions that are detrimental to their well-being.

Measuring Estimation

There have been a number of attempts to quantify estimation abilities following brain injury. Amongst the most common tests of estimation are various forms of cognitive estimation tests (CETs). The original CET was developed by Shallice and Evans (1978) and participants were required to estimate answers to a series of questions where a precise answer was unlikely to be known. The original version was made up of 15 questions and participants were provided with a response sheet and asked to complete the questions with their 'best guesses' in the spaces provided (Strauss, Sherman &

Spreen, 2006). CETs are generally scored according to an error-based scale, which is developed according to a standardised sample, and scored depending on how much answers deviate from the norm. CETs involve providing 'reasonable' answers to questions where relevant knowledge but not the exact answer is available to the individual (Shallice and Evans, 1978). Thus, estimations are taken as analogues of real life judgements. There are many complex cognitive functions involved in cognitive estimation, including activating and accessing semantic memories, working memory, planning, self- monitoring and self-correction (Bullard et al., 2004). Gansler, Varvaris, Swenson and Schretlen (2014) argued that based on the task demands of the CET, this test can be considered a measure of executive functioning and this view appears to predominate throughout the literature. Executive functioning is not considered a unitary process, but has been described by Lezak (1983) as "goal formulation, planning and carrying out goal-directed plans effectively" or as the ability to "organise a sequence of actions towards a goal" (Anderson, Jacobs & Anderson, 2008).

Executive functions, Estimation and the Frontal lobes

Shallice and Evans (1978), who devised the original CET, found that people with anterior lesions performed worse than those with posterior brain lesions. By contrast, other studies (e.g. Taylor & O'Carroll, 1995) have not found this difference. In a functional magnetic resonance imaging (fMRI) study, Horacek et al. (2010) found that the CET activated areas mainly in the frontal lobes, though also with some involvement of the parieto-occipital system.

Validity and Reliability of Cognitive Estimation Tests

For the CET to be clinically useful, its predictive validity in relation to everyday decision-making must be demonstrated. A number of papers have examined the reliability and validity of CETs, but few studies compare it to tasks of everyday functioning. A review by Wagner, MacPherson, Parente and Trentini (2011) aimed to explore the use of CETs in both healthy and neurological populations. They reviewed correlational studies exploring performance on the CET with other measures of executive function and also examined case-control designs. They concluded that the CET has been shown to be associated with frontal lobe injury and executive functions, but that it is currently unclear what particular areas of the frontal lobes are responsible for successful performance. They commented that it is vital that studies exploring this issue consider aetiology and localisation of brain injury when comparing different groups. Their paper did not systematically review the overall quality of the literature, but provided a narrative summary of the results found. Reliability of the CET has also been explored in the literature but to a lesser extent. O'Carroll, Egan and MacKenzie (1994) explored the psychometric properties of the CET and found that it had poor internal reliability but adequate inter-rater reliability (assessed by two authors independently scoring CET responses). Spencer and Johnson-Greene (2009) also assessed the CET's psychometric properties and found that it had limited internal reliability. Macpherson et al. (2014) highlighted concerns over previously developed CETs, such as small control groups, lack of published normative data for the original version by Shallice and Evans (1978), and also questioned its use with those out with the country in which it was developed. They developed two parallel versions of the CET, assessing its reliability, and concluded that it had low internal reliability and, as

found previously, appears to be a relatively multidimensional task, which correlates with multiple cognitive domains (MacPherson et al., 2014).

This Review

No systematic review has examined the validity of CETs. There are a number of different forms of validity, but in essence, *"Validity refers to the degree to which evidence supports the interpretation of test scores for their intended purpose; therefore the examination of a test's validity requires an evaluative judgement by the test user"* WMS-IV Technical and Interpretive Manual (2009).

The existing literature suggests that CETs have been used to identify whether there is damage to the frontal lobes, as a measure of executive functioning, and to determine if an individual would have difficulties making important judgements and decisions in everyday life. The way in which these have been demonstrated is through correlational studies examining the relationship between performance on the CET with other valid neuropsychological tests, or in case-control designs, by comparing differing population groups to assess differences in performance. In relation to case-controls, studies have investigated whether there are differences in CET performance between patients and healthy controls, patient groups with evidence of different anatomical lesions (e.g. frontal vs. posterior lesions) and between patient groups with evidence of executive dysfunction and those without.

In this review, correlational studies were included that explored associations between the CET and other executive functioning tests. In addition, case-control studies comparing a group of patients with a frontal or executive deficit, with a group of

patients with brain injury not specific to the frontal lobes or those deemed not to have executive functioning deficits were included. This will allow comment upon the sensitivity of the CET to functions dependent upon the frontal regions of the brain, and also assess its validity as a test of executive function. Studies only comparing performance of a brain injury group with a healthy control group on the CET were not included in this review since they do not offer any information on the usefulness of the CET in clinical practice nor give any information on the specificity of the CET to any particular brain or cognitive dysfunction. The research evidence, quality of reporting, and quality of methods used to determine the validity of these tests were reviewed.

Method

Search Strategy

The following databases were searched for relevant studies up to 31st March 2015: MEDLINE, EMBASE, PsycINFO and Web of Science. No restrictions were put in place. The search term cognitive estimat* was the only search term used given the limited literature on this area. A requirement for the words to be consecutive was set.

Eligibility Criteria

Inclusion criteria. Studies were included if they were published in English, were peer reviewed publications**,** were case-control studies or correlational studies evaluating the validity of CETs in adult (18+) populations without any developmental disorders. The precise assessments that were included as CETS were those with questionnaire based items, those that were developed based on the original CET, and had 'cognitive estimation/estimate test/task' in its title. If a case-control study was identified, papers were only included where patient groups had clear differences in either the location of anatomical damage (frontal versus non-frontal), or clear differences in their executive functioning identified by specific neuropsychological tests. If a correlational study was identified, it was included if it explored correlations of scores on the CET with scores on other executive functioning tests.

Exclusion criteria. Studies were excluded if no CET outcomes were reported or if the CET was not a standard questionnaire based test. If a case-control study was identified, it was excluded if the only comparison was between patients with brain injury and a healthy control group.

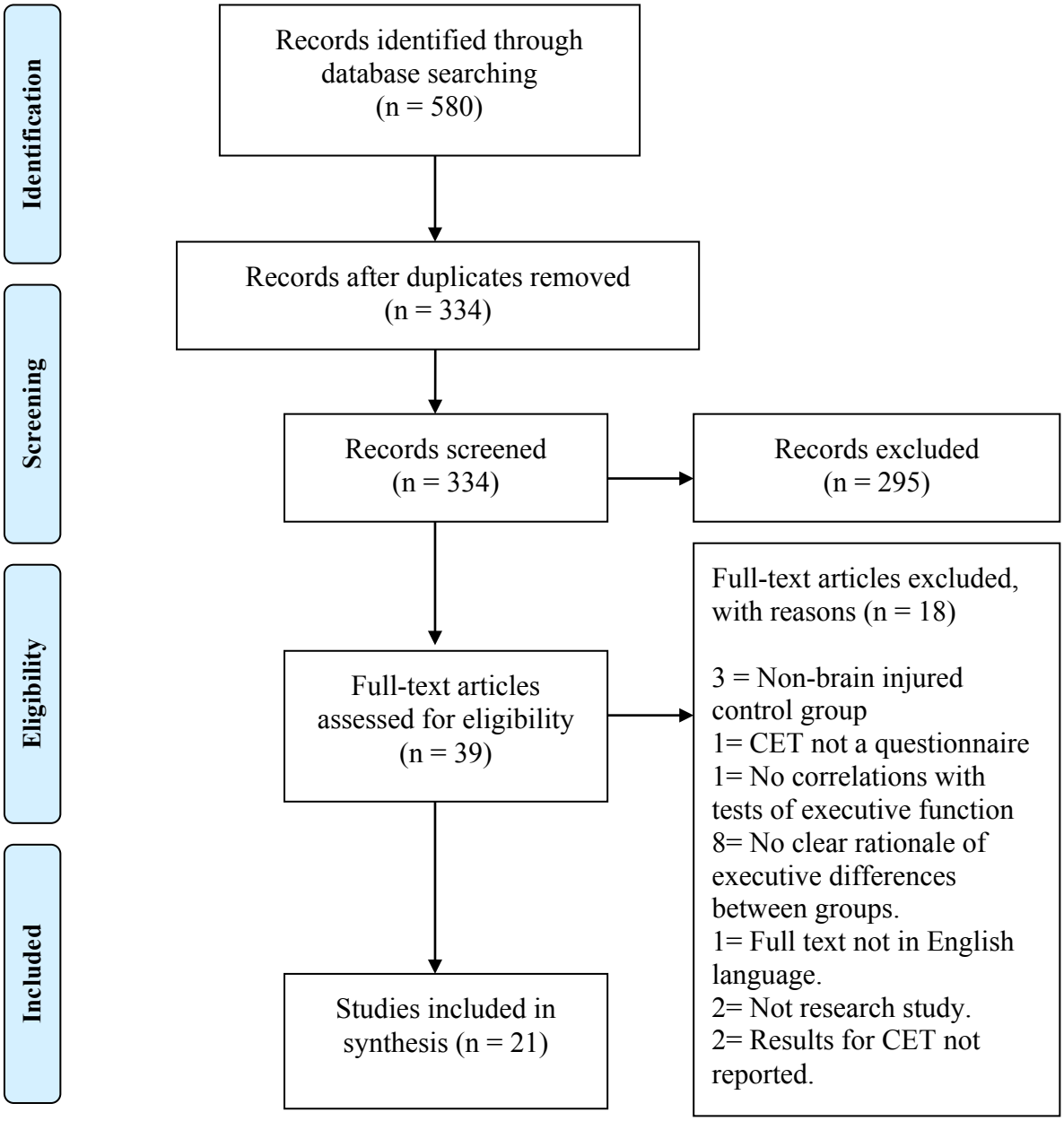
Study Selection and Data Extraction

The main author reviewed all titles and abstracts identified by the search. Articles were selected for full-text review for any titles or abstracts that appeared to satisfy the inclusion criteria, or for which inclusion or exclusion could not be clearly determined at this point in the process.

Search results were compiled using citation management software (RefWorks version 2.0; ProQuest, http://www.refworks.com).

The PRISMA flow diagram of study selection (Moher, Liberati, Tetzlaff, & Altman, 2009) shown below provides a summary of the process used for selection of the studies in this review (see Figure 1).

Figure 1 Flow diagram of study selection



Search Results

The search for studies retrieved 580 unique citations (see Figure 1 for PRISMA flow diagram). Of these, 541, including duplicates were excluded at the title and abstract stage and 39 were examined in full-text. Twenty-one articles met the criteria for inclusion in this review.

Data Extraction. Data from all included studies were extracted by one reviewer (CT) using a form designed specifically for this review. Data extracted included: authors, sample characteristics (age, gender, education, recruitment), type of CET, country in which study was conducted, study design, effect size and main findings. Effect size was calculated for most studies.

Rating of Included Studies. Studies were evaluated according to The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE-22). The STROBE statement is a checklist of 22 items that should be addressed in articles reporting on three study designs of analytical epidemiology, namely cohort, case-controls, and cross-sectional studies (von Elm et al., 2007). An additional eight questions were included relating to the quality of the methodology in each study. These were developed based on methodological issues considered important in relation to this specific topic, and included questions from the SIGN Methodology Checklist for case-control studies (2007). Ratings were represented in terms of percentage scores. See Appendix 2.1 and 2.2 for a list of the items rated.

Papers were initially rated by the main author (CT) and were then blindly co-rated by a PhD student (MJ) carrying out research in brain injury. The STROBE-22 items were rated in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE-22): Explanation and elaboration (2007). The overall

correlation between the scores was r= .564. Mean percentage agreement for each item across all studies on the STROBE-22 was 75%. In 85.7% (n=18) of the studies, total scores on the STROBE-22 showed greater than 1-point of deviation. Discrepancies were discussed and a final score derived for each item. Discrepancies arose mainly due to issues of subjectivity in interpreting Vandenbroucke et al.'s (2007) guidelines for scoring the STROBE-22. The items which most commonly led to disagreement were the following: whether information included in the abstract was informative and balanced; how elements of the study design were presented; how variables were defined; issues around sources of data and details of assessment methods; how quantitative variables were handled in the analyses; and description of statistical methods. For example, one of the questions asked whether a study explained how quantitative variables were handled in the analyses. For this item, one rater scored 0 if this was not explicitly mentioned in the text, however the other rater gave a point for reference to a validation study. It was agreed that standardised neuropsychological tests would likely be reported as continuous variables as long as they were cited with validation studies, and therefore should get a point. If the CET was the main outcome variable then a description and explanation of the scoring system should have been given (or reference to a validation study) in order to get a point for this item. Another example of a source of disagreement was whether papers got a point for describing all statistical methods. Some of the sub-items were not applicable to all study designs. For the sub-item relating to missing data, one rater gave 0 points if missing data was not mentioned in the text, however the other rater gave a point if participant numbers throughout the study were noted in a table format. In most studies, missing data was not addressed in data analysis. This may have been because there was no missing data. Therefore we agreed that if Ns were reported with the main results, or if degrees of

freedom were given which indicated that all the participants who were reported to have taken part, completed all the tests, then the study was given a point. For this and other sub- items, a study was only given a point if they were marked favourably for all relevant item sub-sections.

For the methodological quality questions, mean percentage agreement for each question was 89.8% and in only 9.5% of the studies was there greater than 1-point of deviation. See Appendix 2.3 for rating scores. A third rater (JE) was available should agreement between raters not be possible but this was not required. Due to the heterogeneity of study designs and methods of determining validity, it was not appropriate to perform meta-analysis. Analysis was by narrative synthesis and data were tabulated.

Results

Study Characteristics

Table 1 provides a summary of the included papers. Eight correlational studies, 11 case-control studies and two studies including both types of analyses were identified. Table 2 shows characteristics of CETs that were used in different studies. Table 3 provides a summary of the ratings for the STROBE-22, and Table 4 a summary of methodological quality ratings. Studies were rated according to the particular areas of the studies that were the focus of this review. For example, if a study assessed a patient group versus a healthy control group, but also carried out correlational analyses, we assessed this as a correlational study. Given the differences in outcomes and design, the two different types of studies will be discussed separately with regards to the methodological quality rating. This will ensure that fair comparisons are made and sound conclusions can be drawn about the papers, since three out of eight of the quality questions were not applicable to the correlational studies.

Methodological Quality Rating

The quality of the reporting of the studies ranged from 18.2%- 63.6%. Studies were categorised into low, medium and high quality. These categories were determined a priori. High quality articles were rated as greater than 55% (6 papers); moderate quality as 40-55% (12 papers); and low quality as less than 40% (3 papers). Where effect sizes were not reported in studies, they were calculated if sufficient data were present (see Table 1). For the STROBE-22 items, there were certain items which were reported well in the majority of studies, and some that were poorly reported (see Table 3). For example, no papers reported how the study size was determined. Overall, details of

methods of assessment and a clear summary of results with reference to objectives were well reported across studies, with over 95% of the papers fulfilling these criteria.

For the eight methodological quality questions, studies ranged from 25-100%, showing a wide range of quality across studies. When discussing the papers, the overall quality will be discussed primarily with regard to these methodological quality questions with additional information being provided by the STROBE-22 checklist where appropriate. In order to provide a clear sense of the overall high, moderate and low quality studies, with regards to both reporting and overall quality of methodology, the studies were categorised from A- E (see Table 5).

Table 1 Data Extraction Table

Study	Sample Characteristics (age, gender, education, recruitment)	Type of CET	Country in which study was conducted	Study Design	Effect Size	Main Findings
Appollonio et al. (2003) Cognitive estimation: comparison of two tests in nondemented parkinsonian patients.	**Sample:** 30 patients with idiopathic Parkinson's Disease (PD) without dementia. Defined by MMSE adjusted score <24, UPDRS- III score. **Recruitment:** Not reported. **Mean age:** 66.1 **Gender:** 19M 11F **Education (years):** 6.8	**Italian**	**Italy**	**Correlation**	r= 0.31	Non-significant correlation between the CET and the Frontal Assessment Battery. Therefore in agreement with previous data arguing against a prominent dysexecutive nature of cognitive estimation deficits.
Barrera et al. (2005) Formal thought disorder in schizophrenia: an executive or a semantic deficit?	**Sample:** patients diagnosed with schizophrenia with (n=15) and without (n=16) formal thought disorder (FT), controls (n=17). Patients all met RDC criteria for chronic schizophrenia. Subgroups defined according to global scores on the global rating of Positive Formal Thought Disorder scale. **Recruitment:** Not reported. **Mean age:** Controls 4.1** Non-FT 41.2 FT 47.1 **Gender:** Total patient group 24M 7F, Controls 9M 8F **Education/IQ:** All patient groups WAIS IQ of 85+	**Shallice & Evans (1978)**	**UK**	**Case-control**	r = 0.44	The patients with formal thought disorder were significantly impaired compared to the non-formal thought disordered patients. Formal thought disorder in schizophrenia may be the result of a combination of executive and higher order semantic function. Supports hypothesis that formal thought disorder involves some kind of executive function, and CET distinguishes between the two subgroups.
Brand et al. (2003a) Cognitive estimation and affective judgments in alcoholic Korsakoff patients.	**Sample:** 41 patients with Korsakoff's syndrome (KS) diagnosed according to ICD-10 and DSM-IV criteria, 39 healthy controls. **Recruitment:** KS recruited from 4 different homes for chronically multi-impaired addicts. **Mean age:** KS 56.8, HC 59.7 **Gender:** KS 23M 16F, HC 22M 17F **Education (years):** KS 27(9 or less y) 6(10y) 6(11+y) KS 22(9 or less y) 13(10y) 4(11+y)	**German TKS**	**Germany**	**Correlation**	r= 0.37, 0.51, -0.55	Significant correlations between both total score on CET and bizarre error score, with other tests of executive functioning.

Brand et al. (2003b) Cognitive estimation in patients with probable Alzheimer's disease and alcoholic Korsakoff patients.	**Sample:** 50 patients with probable Alzheimer's disease (AD) according to NINCD-ADRDA criteria exhibiting mild-mod dementia, 50 patients with clinically diagnosed alcoholic Korsakoff's syndrome according to ICD-10 and DSM-IV, 50 healthy controls. **Recruitment:** AD patients from the Clinic of Neurology of the University of Cologne, KS patients from different homes for chronically-multi-impaired addicts. **Mean age:** AD 67.5, KS 56.3, HC 64.8 **Gender:** AD 25M 25F, KS 32M 18F, HC 19M 31F **Education (years):** AD 34(9y) 7(10y) 9(12y) KS 38(9y) 6(10y) 6(12y) HC 21(9y) 16(10y) 13(12y)	**German TKS**	**Germany**	**Case-Control**	r= 0.37	AD patients were more impaired than KS patients. KS patients expected to perform more poorly due to showing deficits in frontal executive function in the past. AD more general cognitive decline. Does not support hypothesis that CET is sensitive to frontal or executive presentations.
Bullard et al. (2004) The Biber Cognitive Estimation Test.	**Sample:** 28 patients with Alzheimer's disease, 24 patients with Parkinson's disease, 25 healthy controls. **Recruitment:** Dementia patients recruited from a private practice and another medical centre. Diagnosed using DSM-IV and NINCDS-ARDA criteria. **Mean age:** AD 75.1, PD 76.7 **Gender:** AD 18M 10F, PD 18M 6F **Education (years):** AD 12.7, PD 13.3	**BCET**	**USA**	**Case-Control**	ns r= 0.11	Both groups were impaired in their cognitive estimation ability. No significant differences in performance between these two groups. CET does not distinguish these two groups, which might be expected given more frontal executive pathology and presentation of PD patients.
Burgess et al. (1998) The ecological validity of tests of executive function.	**Sample:** 92 mixed neurological patients all with independent diagnosis of brain injury dementia, 216 control participants. **Recruitment:** Patients recruited from UK neurological centres. Controls primarily recruited from a group of individuals who had participated in previous collections of population norms. **Mean age:** Patients 38.5, controls 46.1 **Gender:** Not reported. **Education (WAIS-R FSIQ)**: 92.1	**Shallice & Evans (1978)**	**UK**	**Correlation**	ns r= 0.29, 0.18, 0.10 with the DEX	CET did not correlate with any scores on the DEX. Therefore CET does not appear to be related, as might be expected if a useful measure of executive functioning, to overall levels of executive problems n everyday life.

Dixon et al. (2004) Effect of symptoms on executive function in bipolar illness.	**Sample:** 15 manic, 15 depressed, 15 remitted patients with bipolar diagnosed according to DSM-IV (mania and depression assessed using total scores of Young's Mania Rating Scale and the BDI), 30 healthy controls. **Recruitment:** From the South London and Maudsley NHS Trust **Mean age:** Manic: 34.3, Depressed 33.9, Remitted 35.7, HC 35.2 **Gender:** Manic: 7M 8F, Depressed: 6M 9F, Remitted: 8M 7F, HC 17M 13F **Education (years):** Manic 13.0, Depressed 13.9, Remitted 15.6, HC 12.8	**Shallice & Evans (1978)**	**UK**	**Case-Control**	ns Man vs. dep r= 0.28 Man vs. rem r= 0.23 Dep vs. rem r= 0.05	No differences between bipolar groups on the CET but some significant differences between groups on others tests of executive functioning. If CET was a test that was sensitive to frontal/executive functioning deficits, then it might have been hypothesised to distinguish between performances between patients in the manic group (more frontal presentation) than the other two groups.
Kopelman (1991) Frontal dysfunction and memory deficits in the alcoholic Korsakoff syndrome and Alzheimer-type dementia.	**Sample:** 16 patients with Korsakoff's syndrome (KS), 16 patients with Alzheimer's disease (AD), 16 controls. **Recruitment:** KS patients selected from larger group of patients who had attended psychiatric hospitals around London. Controls consisted of elderly healthy people living in sheltered. accommodation in London and non-academic staff from Institute of Psychiatry. **Mean age:** KS 53.7, AD 68.7, HC 61.7 **Gender:** KS 11M 5F, AD 5M 11F, HC 7M 9F **Education (WAIS FSIQ):** KS 105.3, AD 88.6	**10 items from Shallice & Evans (1978)**	**UK**	**Correlation**	ns Ranged from r= -0.1 to 0.05	CET scores did not correlate significantly with performances on any other of the frontal tests in either patient group. In the KS group, a correlation of 0.52 with the modified Weigl score was in the opposite direction from prediction. CET was one of the tests that correlated least with the other tests such as measures of IQ and retrograde and anterograde amnesia which may suggest it may be better conceptualised as measuring some aspect of access to semantic memory. Overall findings confirm that AD patients are severely impaired, and KS patients more moderately impaired at tests of 'frontal' function.

Leng & Parkin (1988) Double dissociation of frontal dysfunction in organic amnesia.	**Sample:** 7 patients with Korsakoff's syndrome (KS) and 5 patients who had become amnesic following an attack of Herpes Simplex encephalitis (PEn), 7 controls. **Recruitment:** Not reported. **Mean age:** PEn 45.6, KS 60, control 52.9 **Gender:** Not reported. **Education (WAIS FSIQ):** PEn 95.4, KS 100.9	**Shallice & Evans (1978)**	**UK**	**Case-control**	r = 0.8	The post encephalitic group performed more poorly than either the KS patients or the controls ($p<0.05$) (opposite result for the WCST- double dissociation). 'Normal' performance on CET by KS patients (no difference from controls).
Levinoff et al. (2006) Cognitive estimation impairment in Alzheimer disease and mild cognitive impairment.	**Sample:** 40 normal elderly controls (NEC), 73 patients with mild cognitive impairment (MCI). **Recruitment:** NEC recruited from various locations such as community volunteers, hospital and university. MCI patients recruited from memory clinic, referred on basis of complaints of memory loss by themselves or family- objective evidence of memory loss on mental status testing, but not impaired enough to meet the NINCDS-ADRDA criteria for probable AD. 40 AD patients who met NINCDS-ADRDA criteria for the diagnosis of probable AD, all mild-mod in dementia severity according to the Washington University Clinical Dementia Rating Scale. **Mean age:** NEC 74.1, MCI 74.0, AD 78.8 **Gender:** Not reported. **Education (years):** NEC 13.8, MCI 12.7, AD 10.7	**Shallice & Evans (1978)**	**Canada**	**Both**	r = 0.56	The AD group performed significantly worse than the MCI group. Negative correlation between the CET and semantic verbal fluency tasks in the AD patients just missed conventional significance. No significant correlations between CET and executive function measures in MCI patients. Significant negative correlation between CET and semantic verbal fluency in NEC. Negative correlations expected as a high CET score denotes impaired performance. On the basis of the pattern of results from AD and NEC, conclude that although the CET is sensitive to aspects of executive function, it is not a pure measure.
Manning et al. (2005). Anterior and non-anterior ruptured aneurysms: Memory and frontal lobe function performance following coiling.	**Sample:** 19 patients with ACoAA, 16 patients with non-anterior aneurysms (middle cerebral artery and posterior communicating artery), 35 controls. **Recruitment:** Patients recruited from pool of patients admitted to Foch Hospital, diagnosed with intracerebral aneurysm and treated by GDC therapy. **Mean age:** All patients 45.0 **Gender:** ACoAA 15M 4F, non-ant 5M 11F **Education (years):** ACoAA 12.7, non-ant 11.6	**Shallice & Evans (1978)**	**France**	**Case control**	ns r= 0.44	No significant differences on the majority of tests sensitive to fronto-temporal dysfunction between anterior and non-anterior groups including the CET. No association between localization of aneurysm and cognitive performance on CET.

Mendez et al. (1998) Use of the cognitive estimations test to discriminate frontotemporal dementia from Alzheimer's disease	**Sample:** 31 FTD, 31 AD patients, 31 elderly controls. **Recruitment**: Patients recruited from the St Paul-Ramsey Neurocognitive laboratory, the neurobehavioural unit of the West Los Angeles Veteran's Affairs Medical Centre, and Neurological clinics at the University of California at LA. Controls recruited from the community or were spouses of patients. **Mean age:** FTD 65.6, AD 74.6, NEC 67.1 **Gender:** FTD 17M 14F, AD 16M 15F, NEC 17M 14F **Education (years):** FTD 14.1, AD 13.7, NEC 14.1	**Modified CET with 16 quantit-ative questions**	**USA**	**Case-control**	r= 0.32	Contrary to expectation, AD patients provided more extreme estimates in comparison with the FTD patients. Remained a significant result even after co-varying for age differences, MMSE score and overall measure of dementia severity. When the range of scores was reviewed, the degree of overlap did not help distinguish FTD from AD patients.
Nedjam et al. (2004) Confabulation, but not executive dysfunction discriminate AD from frontotemporal dementia.	**Sample:** 22 probable AD meeting DSM-IV criteria for dementia and NINDS-ADRDA criteria for probable AD, 10 probable FTD, based on the Lund-Manchester criteria, 32 controls. **Recruitment:** Not reported. **Mean age:** AD 74.9, AD controls 74.5, FTD 60.6, FTD controls 60.6 **Gender:** AD 7M 15F, AD controls 7M 15F, FTD 7M 3F, FTD controls 7M 3F **Education (years):** AD 9.0, AD controls 7.5, FTD 9.2, FTD controls 7.9	**Shallice & Evans (1978)**	**France**	**Case-control**	ns r= 0.08	No significant difference in performance scores between the AD and FTD groups. Both patient groups are impaired on executive functions. Tests considered sensitive to executive function, such as the CET, are not specific enough to discriminate between those with confirmed/supposed frontal lobe pathology (FTD) and those with AD who not have evident frontal pathology. CET may be sensitive to a different brain area.
Parente et al. (2013) Investigating higher-order cognitive functions in temporal lobe epilepsy: cognitive estimation	**Sample:** 108 patients with drug resistant temporal lobe epilepsy (TLE), 51 healthy controls. **Recruitment:** **Mean age:** TLE 38.4, HC 38.6 **Gender:** TLE 52M, 56F, HC 28M 23F **Education (years):** TLE 11.7, HC 13.5	**Shallice & Evans (1978) & Italian normative data**	**Italy**	**Correlation**	Ranged from: r= -0.51 to 0.29	The correlations between total CET scores and the following executive tests (RCPM, Word fluency on phonemic cue) were modest in size. The correlations between bizarreness CET scores and the following executive tests (RCPM, Word fluency on phonemic cue, TOL, TMTB) were also modest in size. These findings suggest that in patients with TLE, altered lexical-semantic and visual attention abilities, but not executive deficits, can accentuate cognitive estimation impairment.

Roth et al. (2012) Performance on the Cognitive Estimation Test in schizophrenia.	**Sample:** 42 patients with schizophrenia, 42 healthy controls (HC). **Recruitment:** Patients diagnosed using SCID-IV-P, HC recruited through advertisements. **Mean age:** Patients 38.9, HC 35.5 **Gender:** Patients 21M 21F, HC 27M 15F **Education/WASI FSIQ:** Significantly lower in the patient group	**CET (Axelrod & Millis, 1994)**	**USA**	**Correlation**	CET with: WCST errors r= 0.33 DKESF TMT r = 0.31	No significant correlations observed between CET and other measures in HC. In the patient group, CET performance was associated with greater impairment on the TMT and WCST perseverative errors. Therefore, the construct validity, specifically convergent validity of the CET was supported by the presence of significant associations between this test and other tests of executive function in the patient sample. Associated with scores on the WCST and TMT, reflecting aspects of executive functioning such as problem-solving ability and cognitive flexibility. It also related to poorer verbal learning, auditory attention and lower intellectual functioning. Therefore evidence of discriminant validity of CET as specific measure of executive function in this patient group is limited.
Silverberg et al. (2007) Cognitive estimation in traumatic brain injury.	**Sample:** 77 patients with TBI. **Recruitment:** From the Southeastern Michigan TBI Systems programme. Had to meet one of the following 3 criteria: 1. Posttraumatic amnesia duration >24hours, 2. Trauma-related to intracranial neuroimaging abnormalities, and 3. Glasgow Coma scale score of less than 13 in the emergency department. **Mean age:** 43.7 **Gender:** 65M 12F **Education (years):** 12.0	**BCET**	**USA**	**Correlation**	Range from r= -0.36 to 0.47	BCET scores correlated moderately with other standard measures of executive functioning, and contrary to hypotheses, at least as high with other neuropsychological tests with minimal demands on executive functioning. Although modest correlations with standard measures of executive functioning (WM, set-shifting and response inhibition), these correlations were strongly attenuated by partialing out the variance associated with the semantic memory (non-executive functioning component of the BCET). BCET scores did not predict concurrent functional status, as measured by the Disability Rating Scale. Therefore poor construct (in terms of both convergent and divergent validity) and poor ecological validity in

						this sample of patients.
Shallice & Evans (1978) The involvement of the frontal lobes in cognitive estimation.	**Sample:** 96 patients with BI (45 ant, 51 post), 25 controls with extra-cerebral lesions. **Recruitment:** Admitted to the National hospital, clinically assessed at the time as having unilateral focal cortical lesion confined to no more than 2 lobes. **Mean age:** Ant 48.9, post 46.8 **Gender:** Ant 18M 27F, post 33M 18F **Education:** Not reported.	**Shallice & Evans (1978)**	**UK**	**Case-control**	M(sd) not reported p<. 025 – p<. 05 Average %extreme responses: ant 20.1% post 14.8%	The anterior group performed significantly worse than the posterior group on the CET. The deficit can be dissociated from one of 'general intelligence' or reasoning as co-varying the results with Raven's Matrices left the effect unaltered.
Shoqeirat et al. (1990) Performance on tests sensitive to frontal lobe lesions by patients with organic amnesia: Leng & Parkin Revisited.	**Sample:** 16 patients with KS, 10 who became amnesic after an attack of PEn, and 5 patients with amnesia resulting from ACoAA, 31 controls. **Recruitment:** Patients matched in terms of score of WMS and Warrington Recognition Memory Test. **Mean age:** KS 56.0, PEn 39.0, ACoAA 36.0 **Gender:** Not reported. **Education (FSIQ):** KS 99, PEn 105, ACoAA 93	**Shallice & Evans (1978)**	**UK**	**Both**	ns r= 0.48	CET did not significantly correlate with any executive functioning measures (2 versions of WCST, FAS). No differences between any patients groups in performance scores on CET.
Spencer & Johnson-Green (2009) The Cognitive Estimation Test (CET): psychometric limitations in neurorehabilitation populations.	**Sample:** 112 patients with various neurological impairments **Recruitment:** Receiving routine neuropsychological examinations as part of their rehabilitation care. **Mean age:** 65.3 **Gender:** 55M 57F **Education (years):** 12.5	**CET (Axelrod & Millis, 1994)**	**USA**	**Correlation**	CET with: RCFT r= -0.46, COWAT r = -0.35	CET was moderately correlated with nearly all cognitive tests examined regardless of their executive demands. Poor divergent validity. Because it appears to be correlated with multiple cognitive domains, may be more appropriately viewed as a measure of global cognition rather than as a test of EF.
Taylor & O'Carroll. Cognitive estimation in neurological disorders.	**Sample:** 370 neuropsychological, neuropsychiatric and neurosurgical patients, 150 controls. **Recruitment:** Most diagnosed according to ICD-9 criteria, KS subjects recruited from psychiatric hospital or supportive hostel accommodation in connection with a drug trial and diagnosed using DSM-III-R. Controls recruited from relatives of patients, local volunteer groups, armed services and	**10-item CET as described by Shoqeirat et al. (1990)**	**UK**	**Case-control**	ns r= 0.05	No significant different was found between those with anterior or posterior lesions. Therefore failed to produce evidence supporting the sensitivity of the CET to anterior cerebral pathology.

	healthy subjects from a general practice community survey. **Mean age:** Patients 46.8, controls 51.5 **Gender:** Patients 197M 173F, controls 83M 67F **Education (years):** Patients 11.5, controls 11.6					
Treitz et al. (2009) Executive deficits in generalized and extrafrontal partial epilepsy: long versus short seizure-free periods.	**Sample:** 35 outpatients with generalised epilepsy and extra-frontal partial epilepsy (divided into seizure free <3m and >3m). 16 healthy controls. **Recruitment:** Not reported. **Mean age:** Seizure free <3m 38.3, seizure free >3m 38.6, HC 38.6 **Gender:** Seizure free <3m 10M 8F, seizure free >3m 8M 9F, HC 7M 9F **Education (years):** Seizure free <3m 7.0, seizure free >3m 6.5, HC 8.0 **IQ**: Seizure free <3m 103.4, seizure free >3m 106.1, HC 112.7	**German TKS**	**Germany**	**Case-control**	ns r= 0.03	No significant differences between groups on CET, but significant difference between groups on phonemic letter fluency. Those who were seizure free <3m generated significantly fewer phonemic items than controls and those <3m produced more overall errors than those >3m seizure free, p= 0.031).

** Reported in paper but likely to be an error

Acronym	Meaning
ACoAA	Anterior Communicating Artery Aneurysm
AD	Alzheimer's Disease
BCET	Biber Cognitive Estimation Test
BDI	Beck Depression Inventory
CET	Cognitive Estimation Test
DEX	Dysexecutive Questionnaire
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th Edition
EF	Executive Functioning
FSIQ	Full Scale Intelligence Quotient
FT	Formal Thought Disorder
FTD	Frontal Temporal Dementia
GDC	Guglielmi Detachable Coils Therapy
HC	Healthy Controls
ICD-9/ICD-10	International Classification of Diseases 9th and 10th Revisions
IQ	Intelligence Quotient
KS	Korsakoff's Syndrome
MCI	Mild Cognitive Impairment
MMSE	The Mini Mental State Examination
NEC	Normal Elderly Controls
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association
PD	Parkinson's Disease
PEn	Herpes Simplex Encephalitis
RCPM	Raven's Coloured Progressive Matrices
RDC	Research Diagnostic Criteria
SCID-IV-P	Structured Clinical Interview for DSM disorders
TBI	Traumatic Brain Injury
TKS	Test zum kognitiven Schätzen
TLE	Temporal Lobe Epilepsy
TMTB	Trail Making Test Part B
TOL	Tower of London Test
UPDRS- III	The Unified Parkinson's Disease Rating Scale Part 3
WAIS-IV	Wechsler Adult Intelligence Scale, 4th Edition
WCST	Wisconsin Card Sorting Test

Table 2 Characteristics of CETs in Included Studies

CET	**Content**	**Sample on which scoring based.**	**Reliability Data** (some data from Strauss, Sherman & Spreen, 2006)	**Administration**	**Example of item that differs from other included tests**
Original CET (Shallice & Evans, 1978)	15 items Numerical and non-numerical 10 item version	25 British neurologically intact individuals. Normative data: 150 healthy individuals from UK.	Reliability data: Internal Consistency: Item total correlations range from - .16 to .57 for the American (Axelrod & Millis, 1994), and the British version (Gillespie et al., 2002). O'Carroll et al. (1994) found that the internal consistency of the British version of the CET was .40 (Cronbach's alpha) and .35 (Guttman split-half reliability coefficient). Ross et al. (1996) examined reliability in American college sample (r=158) and reported that internal consistency was low (Cronbach's alpha= .37). Test-retest Reliability: Ross et al. (1996) retested 44 individuals following about 37.5 days (SD= 17.5). Co-efficient of stability for CET was low (r= .57). On average, slightly better scores obtained at re-test (M=4.7, SD=2.1) than at initial examination (M=5.3, SD=2.3). Suggests a moderate practice effect. O'Carroll et al. (1994) reported that the inter-rater reliability coefficient for a subgroup of 50 subjects was $r = .91$ ($p < .001$), for a group of 50 healthy subjects given British version of CET in which responses scored from 0 (good estimate) to 3 (bizarre estimate), despite rater R.O'C. having a mean score of 6.1 (3.8) vs.	Examiner provides a response sheet with test questions and requests that patients complete questions with "best guesses" in the spaces provided. No time limit.	*What is the best paid job or occupation in Britain today?*

		Age 17-91 Mean education (years) 11.6	rater V.E. having a mean score of 5.1 (3.9) *(t*= 4.04, ***p*** < .01).		
Italian CET Appollonio et al. (2003)	20 items	No data found.	No data found for this 20 item test.	No information provided in paper.	*What is the height of a traffic light?*
German TKS (Brand, Kalbe & Kessler, 2002)	16 items Numerical-assessing size, weight, quality, time. Images shown to participants for a number of the questions. 4 questions from each category.	171 cognitively unimpaired subjects. Mean age= 56.1 Age range (20-71+) 79M 92F	The reliability (Cronbach's alpha) is α = .76. No information about test-retest reliability.	Face to face Answers recorded on paper by test administrator. No time limit. If image is shown, it is only shown for a period of 5 seconds.	*How long is the duration of a morning shower?* *How long is a flight from Frankfurt to New York?* *How many paperclips are in this picture?*
BCET (Bullard et al. 2004)	20 items 5 estimation questions in each of 4 categories time/duration, quantity, weight and distance. Scoring based on percentile data.	113 healthy volunteers. Mean age 37.3 SD 16.1 Range (17-85) Mean education (years) 16.5 50M 68F Ethnicity: Black 4 White 108 Native American 1 Cross-validated with an additional 49 normal volunteers. Mean age 40.3 SD 14 Range 17-78	Internal Consistency Bullard et al. (2004) reported insufficient reliability within healthy controls to assess internal consistency. In patients with dementia, reported Cronbach's alpha og .62 and Guttman split-half of .74. Test-retest reliability information not available.	Asked to read instructions and answer questions on paper with examiner present. Prompted to include units and to provide an answer to every question.	*How long does it take for fresh milk to go sour in the refrigerator?* *How high off a trampoline can a person jump?*

		Ethnicity: 90% white, 4% African American, 2% Native American, 2% Asian, 2% Latino Mean education (years) 13.7 SD 3.1 19M 30F			
Mendez et al. (1998)	16 items Numerical questions	No information on sample used to develop percentile scoring method. Modification of CET originally developed by James Mack, PhD. Questions aimed at American subjects.	No information on reliability data.	Administered face to face and questions read aloud by examiner in addition to simultaneous presentations in written form. Required to guess whenever they failed to provide a spontaneous response.	*How many hairs are there on an average woman's head?* *How long does a house fly live?*
Axelrod & Millis (1994)	10 items Numerical questions. Empirically based standardised scoring method.	164 employed adults recruited from 2 university medical centres. Mean age 39.0 Mean education (years) 16.2 42M 122F Ethnicity: White 123 Black 37 Other 4 Only 143 completed protocols.	No information on reliability data.	Administered face to face and questions read aloud by examiner. Provided units for each item.	*What is the average temperature in Anchorage, Alaska on Christmas Day?*

Acronym	*Meaning*
AD	Alzheimer's Disease
BCET	Biber Cognitive Estimation Test
CET	Cognitive Estimation Test
KS	Korsakoff's Syndrome
MMSE	Mini Mental State Examination
TMT	Trail Making Test
WCST	Wisconsin Card Sorting Test
WAIS	Wechsler Assessment of Intellectual Functioning

Table 3 STROBE-22 Ratings

Strobe Question ***Paper***	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***13***	***14***	***15***	***16***	***17***	***18***	***19***	***20***	***21***	***22***	***T***	***%***	Quality percentage
Appollonio et al. (2003)	***0***	***0***	***0***	***1***	***0***	***0***	***0***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***1***	***0***	***0***	***1***	***0***	***0***	***0***	***0***	***4***	**18.2**	25
Barrera et al. (2005)	***0***	***1***	***1***	***1***	***0***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***11***	**50**	62.5
Brand et al. (2003a)	***0***	***1***	***1***	***0***	***0***	***1***	***1***	***1***	***0***	***0***	***1***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***12***	**54.5**	100
Brand et al. (2003b)	***0***	***1***	***0***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***1***	***0***	***0***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***10***	**45.5**	80
Bullard et al. (2004)	***0***	***1***	***0***	***0***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***0***	***0***	***0***	***9***	**40.9**	87.5
Burgess et al. (1998)	***0***	***1***	***1***	***1***	***0***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***1***	***0***	***1***	***1***	***1***	***1***	***0***	***1***	***12***	**54.5**	80
Dixon et al. (2004)	***0***	***0***	***1***	***0***	***0***	***1***	***1***	***1***	***0***	***0***	***1***	***1***	***1***	***1***	***1***	***1***	***1***	***1***	***1***	***1***	***0***	***0***	***14***	**63.6**	75
Kopelman (1991)	***1***	***1***	***1***	***0***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***1***	***13***	**59.1**	75
Leng & Parkin (1988)	***0***	***1***	***0***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***1***	***0***	***0***	***1***	***0***	***0***	***1***	***0***	***5***	**22.7**	25
Levinoff et al. (2006)	***0***	***1***	***1***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***1***	***1***	***1***	***1***	***1***	***1***	***0***	***1***	***13***	**59.1**	100
Manning et al. (2005)	***0***	***1***	***1***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***1***	***0***	***0***	***0***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***11***	**50**	62.5
Mendez et al. (1998)	***0***	***0***	***1***	***0***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***1***	***0***	***1***	***0***	***1***	***1***	***1***	***0***	***0***	***10***	**45.5**	75
Nedjam et al. (2004)	***0***	***1***	***0***	***0***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***10***	**45.5**	75
Parente et al. (2013)	***0***	***1***	***1***	***0***	***0***	***1***	***1***	***1***	***0***	***0***	***1***	***1***	***0***	***0***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***11***	**50**	80
Roth et al. (2012)	***0***	***1***	***1***	***0***	***0***	***1***	***1***	***1***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***1***	***1***	***1***	***1***	***0***	***14***	**63.6**	80
Silverberg et al. (2007)	***0***	***1***	***1***	***1***	***1***	***1***	***1***	***1***	***1***	***0***	***1***	***1***	***0***	***0***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***14***	**63.6**	60
Shallice & Evans (1978)	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***1***	***0***	***1***	***0***	***0***	***8***	**36.4**	37.5
Shoqeirat et al. (1990)	***0***	***1***	***1***	***0***	***0***	***0***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***1***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***10***	**45.5**	25
Spencer & Johnson-Green (2009)	***0***	***1***	***1***	***0***	***0***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***1***	***1***	***0***	***1***	***13***	**59.1**	80
Taylor & O'Carroll (1995)	***0***	***0***	***1***	***0***	***0***	***1***	***0***	***1***	***0***	***0***	***0***	***1***	***0***	***1***	***1***	***0***	***1***	***1***	***0***	***1***	***0***	***0***	***9***	**40.9**	75
Treitz et al. (2009)	***0***	***1***	***1***	***0***	***0***	***0***	***1***	***1***	***0***	***0***	***1***	***0***	***0***	***1***	***1***	***0***	***1***	***1***	***1***	***1***	***1***	***0***	***12***	**54.5**	62.5
TOTALS	***1***	***17***	***16***	***5***	***5***	***17***	***17***	***20***	***7***	***0***	***8***	***6***	***1***	***11***	***18***	***3***	***17***	***21***	***7***	***18***	***6***	***4***			

Table 4 Methodological Quality Rating

Extra Quality Question ***Paper***	1	2	3	4	5	6	7	8	Quality percentage
Appollonio et al. (2003)									25
Barrera et al. (2005)									62.5
Brand et al. (2003a)									100
Brand et al. (2003b)									100
Bullard et al. (2004)									87.5
Burgess et al. (1998)									80
Dixon et al. (2004)									75
Kopelman (1991)									60
Leng & Parkin (1988)									25
Levinoff et al. (2006)									100
Manning et al. (2005)									62.5
Mendez et al. (1998)									75
Nedjam et al. (2004)									75
Parente et al. (2013)									80
Roth et al. (2012)									80
Silverberg et al. (2007)									60
Shallice & Evans (1978)									37.5
Shoqeirat et al. (1990)									25
Spencer & Johnson-Green (2009)									80
Taylor & O'Carroll (1995)									75
Treitz et al. (2009)									62.5
% criteria met	100	71.4	52.4	28.6	90.5	42.9	85.7	78.6	

Key

Yes *No* *n/a*

Table 5 Summary of Quality Categories

QUALITY Rating	STROBE-22 Rating	Papers	Overall Quality Category
High	High	Roth et al. (2012) Spencer & Johnson-Greene (2009) Levinoff et al. (2006)	A
High	Moderate	Brand et al. (2003a) Brand et al. (2003b) Bullard et al. (2004) Burgess et al. (1998) Parente et al. (2013)	B
Moderate	High	Dixon et al. (2004) Kopelman (1991) Silverberg et al. (2007)	C
Moderate	Moderate	Barrera et al. (2005) Manning et al. (2005) Mendez et al. (1998) Nedjam et al. (2004) Taylor & O'Carroll (1995) Treitz et al. (2009)	D
Low	Low	Appollonio et al. (2003) Leng & Parkin (1988) Shallice & Evans (1978) Shoqeirat et al. (1990)	E

Quality Criteria

The case-control and correlational studies will be discussed separately below.

Case control studies.

High quality case-control studies. Three articles were identified as high quality and showed little evidence for the CET being a specific measure of frontal dysfunction. Levinoff et al. (2006) highlighted the multiple cognitive deficits involved in Alzheimer's disease (AD) and explored the executive functioning deficits often

exhibited at later stages as the disease moves beyond the amnestic phase. As predicted, the AD group were found to perform significantly worse than the MCI group. Bullard et al. (2004) found that both the patients with AD and Parkinson's disease (PD) were impaired in their cognitive estimation ability with no significant differences in performance. If hypotheses had been established based on the CET being specific to frontal deficits, then the expectation would be that PD patients, with a more frontal, executive pathology and presentation, would perform worse. Brand, Kalbe, Fujiwara, Huber and Markowitsch (2003b) argued that CETs make demands not just on frontal-executive functions but also on aspects of memory such as semantic memory. Hence they predicted that patients with Korsakoff's syndrome (KS; likely to have memory and executive impairments) would perform better than AD patients, who are likely to have more general impairments including semantic memory deficits. This is exactly what they found, with AD patients being significantly more impaired than KS patients. Thus they argued that whilst CETs may be sensitive to executive dysfunction, they are not specific, as they will be affected by other cognitive impairments and in particular semantic memory deficits. Overall, the studies do not demonstrate specificity of the CET to frontal dysfunction.

Moderate quality case-control studies. With the exception of Barrera, McKenna and Berrios (2005), who found that patients with formal thought disorder were significantly impaired on the CET compared to the non-formal thought disordered patients, the majority of these studies showed little evidence showing sensitivity or specificity of the CET to frontal or executive dysfunction. Mendez, Doss and Cherrier (1998) even found that AD patients, who are known to have greater temporal-parietal pathology, provided more extreme estimates in comparison with the patients with

frontotemporal dementia (FTD), who are known to have predominant frontotemporal atrophy. Nedjam, Devouche and Dalla Barba (2004) found no significant difference in performance scores between AD and FTD groups and concluded that tests considered sensitive to executive function, such as the CET, are not specific enough to discriminate between those with confirmed frontal lobe pathology and those who do not have evident frontal pathology. Two studies found no differences between patient groups on the CET but did find significant differences between groups on other tests of executive functioning (Dixon, Kravariti, Frith, Murray and McGuire (2004) and Treitz, Daum, Faustmann, and Haase (2009)). Dixon et al. (2004) hypothesised that if the CET were sensitive to executive deficits, then it would distinguish performances between the manic bipolar group (more frontal presentation) and the depressed and remitted bipolar groups. They found no differences between bipolar groups on the CET. Similarly, Treitz et al. (2009) explored the relationship between executive impairment and seizure freedom in patients with generalised epilepsy and extra-frontal partial epilepsy, hypothesising that those who were seizure free for more than 3 months would perform better on the CET than those who were not seizure free. Although finding differences between groups on phonemic letter fluency, they found no significant differences between groups on the CET. Taylor and O'Carroll (1995) and Manning, Pierot and Dufour's (2005) results also indicated no significant differences between those with anterior or posterior lesions, therefore failing to produce evidence supporting the sensitivity and specificity of the CET to anterior cerebral pathology.

Low quality case-control studies. Shallice and Evans (1978) found that an anterior lesion group performed significantly worse than a posterior group on the CET, but the other studies showed contradictory results with one showing no differences

between patient groups on CET performance (Shoqeirat, Mayes, MacDonald & Meudell 1990) and the other showing that similar to Brand et al. (2003b), the group known to have more frontal pathology (KS) performed better than the comparison group who had become amnesic following an attack of Herpes Simplex encephalitis (PEn; Leng & Parkin, 1998).

Methodological issues. There are a number of methodological issues worth highlighting. The high quality studies generally fulfilled all methodological criteria except Bullard et al. (2004), which did not have the power to detect at least a medium effect size (d= 0.5). The moderate quality studies all had an appropriate and clearly focused question; used valid, reliable tests; and clearly defined cases, which differentiated from controls. Only Taylor and O'Carroll (1995) demonstrated sufficient power to be able to detect a medium effect size, and Mendez et al. (1998) was the only study that reported using a statistically valid approach to defining 'bizarre' answers to questions. Three papers failed to report appropriate inclusion and exclusion criteria. The lower quality studies each had a clear and appropriately focused question, however none of the studies compared patients at baseline to establish any similarities or differences; controlled for confounding factors; provided appropriate inclusion/exclusion criteria, nor had sufficient power to detect a medium effect size. Furthermore, no studies provided evidence of a statistically valid approach to defining 'bizarre' answers, the primary outcome measure. The study by Shallice & Evans (1978) triggered many similar studies. It should, like many of the other studies, be interpreted with caution, given many factors may have affected the results, including the small sample size, lack of clearly defined outcome measures (e.g. scoring method) and lack of control for potential confounding factors. The variability in methodology makes it difficult to draw

firm conclusions regarding which studies may give the most useful information about the CET.

Correlational studies.

High quality correlational studies. Five correlational studies were rated as high quality. Three studies found significant correlations between the CET and tests of executive function. Brand et al. (2003a) for example, found correlations between both the total score and bizarre error score on the CET with tests of executive functioning, and Roth, Pixley, Kruck and Garlinghouse (2012) found that CET performance correlated with two executive functioning tests in the schizophrenia group. These correlations may reflect aspects of executive functioning such as problem-solving ability and cognitive flexibility, providing evidence for convergent validity. However, CET performance was also related to poor verbal learning, auditory attention and lower intellectual functioning. Similar findings were reported by Spencer and Johnson-Greene (2005) who explored the psychometric properties of the CET and found that the CET was moderately correlated with nearly all cognitive tests regardless of executive demands. The authors concluded that due to its correlation with multiple cognitive domains, it might be more appropriately viewed as a measure of global cognition rather than as a test of executive functioning. Thus, these studies provided evidence of good convergent validity but poor discriminant validity of the CET. Levinoff et al. (2006) and Parente, Manfredi, Villani, Franceschetti and Giovagnoli (2013) however, found no significant correlations between CET and tests of executive functioning. Parente et al. (2013) however, also found no relationships between the CET and other neuropsychological tests. These papers failed to support the convergent validity of CETs. Burgess, Alderman, Evans, Emslie and Wilson (1998) explored correlations

between the CET and the Dysexecutive questionnaire (DEX), a measure of everyday functioning, finding no correlations between the CET and the DEX. They concluded that the CET does not appear to be related to overall levels of executive problems in everyday life.

Moderate quality correlational studies. Silverberg, Hanks and McKay (2007) found that the Biber Cognitive Estimation Test (BCET) scores correlated moderately with standard measures of executive functioning, but as high with other neuropsychological tests with minimal demands on executive functioning. Importantly however, correlations with measures of executive functioning, such as those assessing working memory, set- shifting and response inhibition, were diminished by partialing out variance associated with the semantic memory (non-executive functioning) component of the BCET. Like Burgess et al. (1998), they also found that scores on the BCET did not predict concurrent functional status, as measured by the Disability Rating Scale. Kopelman (1991) showed that CET scores did not significantly correlate with performances on any other frontal tests in a KS patient group. A significant correlation was found between a modified CET score and the Modified Weigl test, in the opposite direction to what was predicted. Therefore, these studies demonstrated poor construct and ecological validity of the CET.

Low quality correlational studies. Both Appollonio et al. (2003) and Shoqeirat et al. (1990) found no correlations between the CET and tests of executive functioning.

Overall, the higher quality papers demonstrated good convergent validity of the CET, but poor divergent validity. In the Burgess et al. (1998) study, where a wide range of

patients with a range of neurological deficits were recruited, authors concluded that the CET also has poor ecological validity, and therefore could be limited in its ability to produce useful clinical information. In the moderate-low quality studies, evidence showed poor construct validity (in terms of both convergent and divergent validity) and poor ecological validity of the CET. It was suggested that the CET may rely on multiple domains of cognitive functioning (Spencer and Johnson-Greene, 2007); might be better conceptualised as measuring some aspect of access to semantic memory; or reflect pathology at a site other than the frontal lobes (Kopelman, 1991).

Discussion

This is the first review to systematically evaluate studies that have explored the validity of CETs. Overall, it is probable that the reporting and methodology of study designs may not be conducive to providing clear conclusions about the validity of this test as a useful clinical measure. Given that no studies gained higher than 14 out of 22 points, this would indicate overall poor quality of reporting in these studies, which may affect interpreting the outcomes of studies. Nevertheless, regardless of the quality of the study, the results question the validity of current forms of the CET in being able to provide useful clinical information about frontal pathology, executive dysfunction or everyday decision-making. Construct and ecological validity of the CETs were poor across studies. It would be difficult therefore to conclude that this test is a valid measure of what it is intended to assess. Given the contradictory findings, there continues to be much to learn about the CET and it is highly possible that it may draw upon a number of cognitive functions, frontal and non-frontal. There is a clear need for more rigorously reported, higher quality studies, in order to have more confidence in the conclusions made.

In this review, evidence for the validity of the CET was firstly explored in terms of being able to distinguish between patient groups with frontal pathology or clear executive functioning deficits, and groups that did not have specific frontal or executive difficulties. The majority of the papers failed to clearly distinguish between patient groups and some reported results in the opposite direction to that predicted. It was often concluded that the CET may not be specific to frontal or executive functions, but rather may reflect other cognitive functions. The frontal lobes have connections with almost all other areas of the brain (Wagner et al., 2011) involving many complex interactions,

including the frontal lobe, limbic system and posterior cortex (Slattery, Garvey & Swedo, 2001). It may be difficult, therefore, to draw clear conclusions about the specificity of the CET to the frontal lobes or to executive dysfunction, as it might be expected that other brain areas would also affect performance on the CET.

Secondly, the validity of the CET was explored in terms of how similar it is to other valid measures of executive functioning, including those used to assess everyday functioning and decision-making. In general, evidence suggested that when the CET was shown to correlate with executive functioning measures, it also correlated with non-executive neuropsychological measures. In six of the ten studies including correlational analyses, the CET did not correlate significantly with measures of executive function, including two studies exploring associations between the CET and tests measuring everyday functional abilities. This indicates that the CET may not bear any relevance to everyday functioning, which calls into question its clinical utility, if these results are accurate. However, more studies of this nature are warranted.

Main Limitations of Included Studies

Although the inclusion criteria for case-control studies involved patient groups who had clear differences in either the location of anatomical damage or differences in their executive functioning, this was less clear in practice when reviewing studies. Some studies were included based on authors' hypotheses regarding groups having more frontal pathology or executive presentations, and this was necessary due to the complexity of the studies and patient populations, as well as the lack of certainty in the literature regarding areas of brain pathology.

Only three case-control studies had sufficient power to detect a medium effect size and four correlational studies fulfilled this criterion. None of the papers reported statistical analyses explaining how study size was determined. The implications of not having sufficient power in a study are significant, because studies (e.g. Bullard et al., 2004) may have concluded false negative results based on sample sizes not having enough power to detect differences, if they existed. This means that those studies that did not have sufficient power, and did not find significant differences between groups (Bullard et al., 2004, Dixon et al., 2004, Manning et al., 2005, Nedjam et al., 2004, Shoqeirat et al., 1990, Treitz et al., 2009); and those finding no significant associations between the CET and other tests of executive functioning (Appollonio et al., 2003); may not have been able to detect differences if they had been present. Therefore, this makes it difficult to draw clear conclusions about these studies.

Similarly, less than half of the studies reported a statistically valid approach to defining 'bizarre' answers to questions on the CET, which is the primary outcome measure of this test. As specified in this methodological quality item, an approach that is statistically valid would need to have answers that are compared to a sample of healthy individuals or normed on a healthy group, and a cut-off for impaired performance suggested based on the variability of answers. Failing to do this means that differentiating populations based on this outcome may be invalid. It may be that this was defined, but not reported. This emphasises the importance of clear and explicit reporting.

Another difficulty with this body of literature is that due to the complexity of deficits in the patient populations, there were often opposing hypotheses between papers

examining similar patient groups, depending on the comparison group they were using. The lack of clarity on what areas of the brain the CET is reliant upon meant that authors also often provided different rationales and hypotheses for the groups included in their studies. For example, in some studies, patients with KS, deemed to have frontal pathology, were hypothesised to perform more poorly against patients with more general cognitive decline (Brand et al., 2003b); whereas Shoqeirat et al. (1990) hypothesised that KS patients would perform in the 'normal range', or better than the comparison groups. Levinoff et al. (2006) explored executive deficits in AD known to occur in the later stages of this disease, whereas Brand et al. (2003b) argued that those with AD would be more impaired due to more general, cognitive deficits, rather than executive dysfunction specifically. Therefore, although studies provide a rationale regarding differences between groups on executive functioning, synthesis of results across multiple studies is complex and demonstrates the difficulty of drawing clear conclusions.

Finally, the majority of the papers included in this review did not specifically set out to determine the validity of the CET. Therefore, this would have impacted on how they designed the study, which would not have necessarily leant its design or choice of patient population to the aim of determining the validity of the CET specifically.

Strengths and Limitations of the Current Review

The systematic search strategy; the clear, rigorous method of determining inter-rater reliability; and the explicit description of how validity was assessed and reasons for doing this, were strengths of this review. A limitation includes the fact that only studies with specified comparisons of the CET with executive functioning tests were included.

In addition, within the included studies, multiple versions of the CET were assessed. The heterogeneous nature of the CETs is demonstrated in Table 2 by the observation of different types of estimation questions, varying reliability, varying information about the sample the data was normed upon and in the administration of items. This leads to further difficulties comparing these papers, as it becomes unclear exactly what the tests are measuring and to what extent they are measuring the same constructs. For example, some versions of the CET have both categorical and numerical questions and this makes it more difficult to develop clear scoring systems. Unlike the others, Axelrod and Millis (1994) provided units for the answers to those participating in the test, which could mean they are factoring out a cognitive skill that may be vital to the CET. In addition, some of the questionnaires include questions to which some individuals may know the answer, for example, '*how fast does a commercial jet fly'*. Overall therefore, being able to make fair comparisons and draw meaningful overall conclusions about CETs was difficult, given this heterogeneity of form.

Future Research

The CET is a relatively widely used test of executive function and despite the findings in this review, it would be beneficial to continue investigating CETs. Recently, there has been more interest regarding how these questionnaires are developed, and different versions are now being used which attempt to address previous test limitations. Further research might explore these better-developed, newer tests within the context of establishing their use and validity in a scientifically rigorous manner. Future research should adopt high quality methodological designs, including sufficient power analyses and clear, statistically valid outcome measures. It will also be vital for studies to ensure they control for any potential confounding factors, as failure to do this can lead to

inaccurate interpretations of results. Populations included should also have a clear rationale, taking previous research findings into account and providing a clear hypothesis regarding performance on the CET. A recommendation for studies including patients with AD, for example, might be to specify stage of disease, especially if executive deficits are known only to occur at a later stage. It will also be important to develop new CETs with well-defined and developed questions that are less sensitive to aspects of general knowledge and therefore more specific to executive functioning difficulties. For example, including questions that have no exact answer and rely on knowledge most people are likely to possess. It is recognised that planning future research in this area will be challenging, but it is hoped that this will lead to conclusive findings on the use of this test.

Conclusions

It is likely that current CETs are not useful tools in clinical practice for identifying specific deficits in executive functioning, or in relating performance to how an individual may perform tasks or make decisions pertinent to daily living. It may be that the CET draws upon a number of cognitive functions, including those controlled by the frontal lobes, but not specific to this area. However, the limited number of high quality studies makes it difficult to draw firm conclusions, or generalise findings to other patient populations. Mixed results, and a wide array of patient populations being examined, make it difficult to fully comment on the validity of this test. There is a need for more rigorously reported studies, as well as higher quality methodologies in studies, particularly with regards to statistical power, ensuring statistically sound methods to measure outcomes, and clearer rationales for selection of patient groups, in order to aid clarity on the validity of this test. Imaging studies alongside these designs may be beneficial for this purpose.

References

Anderson, V., Jacobs, R., & Anderson, P.J. (2008). *Executive functions and the frontal lobes. A lifespan perspective.* New York, NY: Taylor & Francis.

Appollonio, I.M., Russo, A., Isella, V., Forapani, E., Villa, M.L., Piolti, R., & Frattola, L. (2003). Cognitive estimation: Comparison of two tests in nondemented parkinsonian patients. *Neurological Sciences*, **24**(3), 153-154. doi: 10.1007/s10072-003-0105-3

Barrera, A., McKenna, P.J., & Berrios, G.E. (2005). Formal thought disorder in schizophrenia: an executive or a semantic deficit? *Psychological Medicine*, **35**(1), 121-132. doi: http://dx.doi.org/10.1017/S003329170400279X

Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition and Emotion*, **24**(4), 561-595. doi: 10.1080/02699930903132496

Brand, M., Fujiwara, E., Kalbe, E, Steingass, H-P., Kessler, J., & Markowitsch, H.J. (2003a). Cognitive estimation and affective judgements in alcoholic Korsakoff's patients, *Journal of Clinical and Experimental Neuropsychology*, **25**(3), 324-334. doi: http://dx.doi.org/10.1076/jcen.25.3.324.13802

Brand, M., Kalbe, E., Fujiwara, E., Huber, M., & Markowitsch, H.J. (2003b). Cognitive estimation in patients with probable Alzheimer's disease and alcoholic Korsakoff patients. *Neuropsychologica*, **41**(5), 575-584. Retrieved from http://www.sciencedirect.com/science/article/pii/S0028393202001835#

Brand, M., Kalbe, E., Kessler, J. (2002). Test zum kognitiven Schätzen (TKS). Göttingen: Hogrefe. Retrieved from http://www.testzentrale.de/programm/media/catalog/Test/0423701_p.pdf

Bullard, S.E., Fein, D., Gleeson, M.K., Tischer, N., Mapou, R.L., & Kaplan, E. (2004). The Biber Cognitive Estimation Test. *Archives of Clinical Neuropsychology*, **19**(6), 835-846. doi: 10.1016/j.acn.2003.12.002

Burgess, P.W., Alderman, N., Evans, J., Emslie, H., & Wilson, B.A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, **4**(6), 547-558. doi: 10.1017/S1355617798466037

Dixon, T., Kravariti, E., Frith, C., Murray, R.M., & McGuire, P.K. (2004). Effect of symptoms on executive function in bipolar illness. *Psychological Medicine*, **34**(5), 811-821. doi: 10.1017/S0033291703001570

Gansler, D.A., Varvaris, M., Swenson, L., & Schretlen, D.J. (2014). Cognitive estimation and its assessment. *Journal of Clinical and Experimental Neuropsychology*, **36**(6), 559-568. doi: 10.1080/13803395.2014.915933

Gillespie, D.C., Evans, R.I., & Gardener, E.A. (2002). Performance of Older Adults on Tests of Cognitive Estimation. Journal of Clinical and Experimental Neuropsychology, 24(3). doi:10.1076/jcen.24.3.286.988

Horacek, J., Preiss, M., Tintera, J., Laing, H., Kopecek, M., Spaniel, F., Brunovsky, M., et al. (2010). A Functional Magnetic Resonance Imaging Study of the Cognitive Estimation. *Activitas Nervosa Superior Redeviva*, **52**(3), 187-192. Retrieved from: http://www.hoschl.cz/files/5115_cz_Horacek%20Laing%20activitas.pdf

Kopelman, M.D. (1991). Frontal dysfunction and memory deficits in the alcoholic Korsakoff's syndrome and Alzheimer-type dementia. *Brain*, **114**, 117-137. doi: http://dx.doi.org/ 117-137

Leng, N.R.C., & Parkin, A.J. (1988). Double dissociation of frontal dysfunction in organic amnesia. *British Journal of Clinical Psychology*, **27**(4), 359-362. doi: 10.1111/j.2044-8260.1988.tb00800.x

Levinoff, E.J., Verret, L., Akerib, V., Phillips, N.A., Babins, L., Kelner, N., & Chertkow, H. (2006). Cognitive estimation impairment in Alzheimer's disease and mild cognitive impairment. *Neuropsychology*, **20**(1), 123-132. doi: 10.1037/0894-4105.20.1.123

Lezak, M.D. (1983). *Neuropsychological assessment* (2nd ed). New York: Oxford University Press.

MacPherson, S.E., Wagner, G.P., Murphy, P., Bozalli, M., Cipolotti, L., et al. (2014). Bringing the Cognitive Estimation Task into the 21st Century: Normative Data on Two New Parallel Forms. *PloS ONE*, **9**(7). doi: 10.1371/journal.pone.0092554

Manning, L. Pierot, L., & Dufour, A. (2005). Anterior and non-anterior ruptured aneurysms: Memory and frontal lobe function performance following coiling. *European Journal of Neurology*, **12**(6), 466-474. doi: 10.1111/j.1468-1331.2005.01012.x

Mendez, M.F., Doss, R.C., & Cherrier, M.M. (1998). Use of the cognitive estimations test to discriminate frontotemporal dementia from Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, **11**(1), 2-6. doi: 10.1177/089198879801100102

Moher, D., Liberati, A., Tetzlaff, J., Altman., DG., the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Annals of Internal Medicine,* **151**(4), 264-269. doi: http://dx.doi.org/10.1136/bmj.b2535

Nedjam, Z., Devouche, E., & Dalla Barba, G. (2004). Confabulation, but not executive dysfunction discriminate AD from frontotemporal dementia. *European Journal of Neurology*, **11**(11), 728-733. doi: 10.1111/j.1468-1331.2004.00981.x

O'Carroll, R., Egan, V., & MacKenzie, D.M. (1994). Assessing cognitive estimation. *British Journal of Clinical Psychology*, **33**(2), 193-197. doi: 10.1111/j.2044-8260.1994.tb01110.x

Parente, A., Manfredi, V., Villani, F., Franceschetti, S., & Giovagnoli, A.R. (2013). Investigating higher-order cognitive functions in temporal lobe epilepsy: Cognitive estimation. *Epilepsy & Behavior*, **29**(2), 330-336. doi: 10.1016/j.yebeh.2013.07.031

Ross, T.P., Hanks, R.A., Kotasek, R.S., & Whitman, R.D. (1996). The reliability and validity of a modified Cognitive Estimation Test. Paper presented to the International Neuropsychological Society, Chicago.

Roth, R.M., Pixley, H.S., Kruck, C.L., & Garlinghouse, M.A. (2012). Performance on the cognitive estimation test in schizophrenia. *Applied Neuropsychology: Adult*, **19**(2), 141-146. doi: 10.1080/09084282.2011.595461

Scottish Intercollegiate Guidelines Network (2007). *SIGN 50: A guideline developer's handbook.* Edinburgh: SIGN.

Silverberg, N.D., Hanks, R.A., & McKay, C. (2007). Cognitive estimation in traumatic brain injury. *Journal of the International Neuropsychological Society*, **13**(5), 898-902. doi: http://dx.doi.org/10.1017/S1355617707071135

Shallice, T., & Evans, M.E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex*, **14**(2), 294-303. doi: 10.1016/S0010-9452(78)80055-0

Shoqeirat, M.A., Mayes, A., MacDonald, C., & Meudell, ,P. (1990). Performance on tests sensitive to frontal lobe lesions by patients with organic amnesia: Leng & Parkin revisited. *British Journal of Clinical Psychology*, **29**(4), 401-408.

Slattery, M., Garvey, M., & Swedo, S. (2001). Frontal-subcortical circuits: a functional development approach. In: Lichter DG, Cummings, JL, editors. Frontal-subcortical circuits in psychiatric and neurological disorders. New York: Guilford: 2001, p. 314-33.

Spencer, R.J. & Johnson-Greene, D. (2009). The Cognitive Estimation Test (CET): Psychometric limitations in neurorehabilitation populations. *Journal of Clinical and Experimental Neuropsychology*, **31**(3), 373-377. doi: 10.1080/13803390802206398

Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary (3rd Ed.)*. Oxford, New York: Oxford University Press.

Taylor, R., & O'Carroll, R. (1995). Cognitive estimation in neurological disorders. *British Journal of Clinical Psychology*, **34**(2), 223-228. doi: 10.1111/j.2044-8260.1995.tb01456.x

Treitz, F.H., Daum, I., Faustmann, P.M., & Haase, C.G. (2009). Executive deficits in generalized and extrafrontal partial epilepsy: Long versus short seizure-free periods. *Epilepsy & Behavior*, **14**(1), 66-70. doi: 10.1016/j.yebeh.2008.08.005

Wagner, G.P., MacPherson, S.E., Parente, M.A. & Trentini, C.M. (2011). Cognitive estimation abilities in healthy and clinical populations: the use of the Cognitive Estimation Test. *Neurological Sciences*, **32**(2), 203-210. doi: 10.1007/s10072-010-0449-4

Weschler, D. (2009). *WMS-IV Technical and Interpretative Manual- Fourth Ed*. San Antonio: Pearson.

Vandenbroucke, J.P., Von Elm, E., Altman, D.G., Gøtzsche, P.C., Mulrow, C.D., Pocock, S.J., Poole, C., et al. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Annals of Internal Medicine*, **147**(8), W-163-W-194. doi: http://dx.doi.org/10.1136/bmj.39335.541782.AD

Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., & Vandenbrouke, J.P. (2008). STROBE initiative the strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Preventative Medicine,* **45**(4), 247-251. doi:10.1016/j.ypmed.2007.08.012

Chapter Two: Major Research Project

Developing a Culture Fair Cognitive Estimation Test

Author: Cathy Tran [1]

[1] Mental Health and Wellbeing, University of Glasgow

Correspondence Address:
Mental Health and Wellbeing University of Glasgow
Gartnaval Royal Hospital
1055 Great Western Road Glasgow
G12 0XH
Email: c.tran.1@research.gla.ac.uk

Prepared in accordance with submission requirements for Journal of the International Neuropsychological Society *(See Appendix 1)*

Word count: 8264 *(not including Abstract, Tables, Figures or References)*

Plain English Summary

Developing a Culture Fair Cognitive Estimation Test

Background

We all make judgements in our daily lives. Often, when people acquire traumas to the brain, their judgement and decision-making becomes impaired. A number of measures aimed at assessing judgement have been developed. The most common of these are Cognitive Estimation Tests (CETs), where people are asked questions to which exact answers are not expected, but estimates are required. There are many brain functions involved in decision-making and therefore, many forms of brain injury will cause impairment in this ability. Previous CETs include questions that are not culture fair, being specific to a particular cultural context. For example, one test included a question about the average time of a dental check-up. In highly deprived areas or in cultures where easily available dental care is not the norm, this question would be unfair.

Aims and Questions

To develop a new CET which can be used in all cultures. Questions that were likely to be culturally fair, with which most people would be familiar, included those referring to the physical world, the human body, and

cultural practices known to all. Firstly, answers were gathered from non-brain injured volunteers, to establish a normal range of answers. Tests were then conducted to see whether the CET distinguishes between those with brain injury and those without. A further exploration looked into whether it distinguishes between those with brain injury who are deemed by clinicians to have capacity to make welfare decisions and those who are not. The association between the CET and a test of daily functioning was explored, as well as examining performance over two time points.

Methods

Participants: Individuals without brain injury, a sample of patients with severe brain injury and a comparison group of healthy participants. All participants were able to consent to participate.

Recruitment: Brain injured participants were recruited from two inpatient brain injury rehabilitation centres in the UK. The clinical team referred into the project and potential participants were given information about the study. The patient sample was divided into two groups regarding their capacity to make important life decisions based on psychiatric assessment. This determined whether the measure differentiates those who are able or not able to make important decisions about their lives.

Design: Comparison of groups and exploring relationships between scores on two tests.

*Data collection: V*ia an online questionnaire, by post, and face-to-face testing.

Main Findings and Conclusions

Patients with brain injury performed significantly worse on the test than a matched control group. There were no differences in performance on those deemed able to make important welfare decisions and those who were not. CET performance was not associated with performance on another test of daily functioning in these patients and there was poor association between scores over two time points. Based on results from this study, CETs do not appear to be useful for clinical purposes. Future studies should ensure that the sample of healthy participants is made up of a larger range of individuals in terms of education and social economic status, examine its use in other clinical samples, and further explore how it might relate to individuals' abilities to make decisions in their everyday lives.

Scientific Abstract

Objective: Cognitive Estimation Tests (CETs) are used to assess decision-making. Previous versions include culturally- biased questions likely to disadvantage certain sections of the population. This study aimed to develop a new culture fair questionnaire and assess its reliability and validity.

Method: A 30-item questionnaire was developed and assessed for culture fairness. A normative range of answers was gathered, and a scale developed to define level of deviation from typical responses. Performance in a group of people with brain injury was compared to a matched group of healthy controls. Those with brain injury deemed able to make significant life decisions were compared with a group considered to lack this capacity, to determine whether this test may be useful when assessing decision-making capacity. Correlational analyses were conducted to determine whether there was a relationship between the test and performance on the Dysexecutive Questionnaire (DEX), a measure of everyday executive functioning. Test-retest reliability was examined with 30 of the normative sample.

Results: Results confirm previous literature showing that those with brain injury perform significantly worse than healthy controls. The test did not discriminate between patients with and without capacity to make important decisions, did not significantly correlate with the total score on the DEX and demonstrated relatively poor consistency.

Conclusions: Based on these results, CETs do not appear to be reliable or valid enough for use in clinical assessments. A sub-set of the most sensitive items may prove useful, but further work is required to examine the reliability and validity of this item subset in new samples.

Mesh Terms: brain injury, decision-making, judgement, culture, validity, reliability

Introduction

Judgement, or the ability to make considered decisions and come to sensible conclusions, is necessary for effective independent living. Judgement is the result of a process by which evidence is evaluated, chances of different outcomes assessed, and an action decided (Blanchette & Richards, 2010). The process of estimation is considered to draw upon similar processes. To estimate requires the ability draw upon existing knowledge, use that information to generate possible answers (estimates), weigh up the possibilities and select the best option. Estimation therefore may be considered a relatively specific judgement task. Estimation is seen to underpin decision-making, where individuals choose something by drawing on knowledge from multiple sources, and select or avoid options that carry unfavourable outcomes (Blanchette & Richards, 2010). These abilities are central to being able to make reasoned decisions in everyday life.

Estimation is considered to be an executive functioning skill. Executive dysfunction and associated impairments in estimation and judgement make a major contribution to neurobehavioural disability following acquired brain injury (Wood, 2001). The processes involved in estimation are thought to include: identifying the problem or relevant knowledge set; retrieving relevant facts and information; holding and maintaining the problem in working memory; carrying out appropriate manipulations on relevant knowledge; developing an initial estimate; iterative comparison and change against knowledge of the world to judge its reasonableness; repeating any part of the sequence to produce a better response; and then finalising the estimate (Bullard et al., 2004).

Impairment of estimation can have serious implications for the individual and are important to assess. Assessment of a person's ability to estimate may be useful in determining whether someone has the capacity to make important decisions, such as managing finances, regulating behaviour in relationships and making decisions regarding future employment. There have been a number of attempts to quantify estimation abilities following brain injury (BI). The most common of these are cognitive estimation tests (CETs). In these, people are asked to estimate answers to a series of questions where the exact answer is unlikely to be known. Thus estimations are taken as analogues of real life judgements. There are many cognitive functions involved in estimation and therefore, many types of brain dysfunction will cause errors in these questions (Bullard et al., 2004).

Issues with Previous CETs

Previous CETs include Shallice and Evans' (1978) original CET and questions within the Behavioural Assessment of Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emslie, & Evans, 1996). Shallice and Evans (1978) demonstrated sensitivity of the CET to lesions of the frontal lobes. Both tests are, however, culture-specific, with performance being influenced by specific prior knowledge, thus limiting populations for whom the questions are valid. For example, Shallice and Evans' test was developed in London and contained the question "how high is the Post Office Tower?", a question bound by cultural knowledge of the London skyline. The BADS (Wilson et al., 1996) includes cognitive estimation questions that are likely to be easier for people with experience of the content of the questions, for example people who own a dog (for the question 'how long do most dogs live for'?), or people who go the dentist regularly (for the question 'How long does a routine dental check-up take?). It is conceivable that

people from areas with economic deprivation, or who belong to cultures where pet keeping, or routine dental care is not the norm may be disadvantaged.

Della Sala, MacPherson, Phillips, Sacco, and Spinnler (2003) developed a CET standardised on 175 healthy individuals in Italy. Their scale indicated a correct range of answers and ranges to indicate whether the response represented a bizarre response, one that is statistically deviant from the norm. The equation of bizarreness ratings and impairment of judgement was a useful development. Although this measure has many questions with good face validity and apparent culture fairness, some items appear highly culture bound, for example "Approximately how many coffees can a barman in a motorway restaurant make in one hour during rush hour?" Whilst there are some positive elements to this test, there remain problems with culture fairness.

It is proposed that culture fair questions should refer to the physical world, the human body, and cultural practices that have become global and familiar to most people. Questions in a culture fair test would place less reliance on recall of cultural information with more emphasis on immediately knowable aspects of the body and the physical world.

Aims

To develop a set of culture fair cognitive estimation questions and assess the reliability and validity of this measure. To collect data from a normative sample and develop a method of scoring; and to test the questionnaire on a group of persons with brain injury and a matched control group to demonstrate whether this test distinguishes between these groups. To further explore performance in those with brain injury who are considered to have difficulties making decisions regarding important aspects of everyday living compared with those who are not.

Hypotheses

Primary hypothesis. People with brain injury will give answers on the Culture Fair Assessment of Cognitive Estimation (CFACE) that significantly differ from those of healthy controls.

Secondary exploratory hypotheses. Those with brain injury deemed by clinicians as unable to make major decisions regarding their welfare, for example, managing household budgets or deciding where to live, will perform at a significantly poorer level on the test than those with capacity to make these decisions.

For participants with brain injury, performance on the CFACE will correlate significantly with ratings of executive functioning (The Dysexecutive Questionnaire, DEX; Burgess, Alderman, Wilson, Evans & Emslie, 1996) and in particular with the executive cognitive sub-scale of this measure.

Methods

Participants

Inclusion criteria normative sample. English speaking male and female volunteers who have given consent to participate.

Exclusion criteria normative sample. Any prior history of psychiatric treatment suggesting disorders with potential neurobehavioural effects; any previous head-injury requiring medical treatment; previous episode of unconsciousness, alcoholism, evidence of any neurological disorder (including stroke, seizures, tumours), chronic medical conditions which might affect neuropsychological function (such as cardiovascular disease, diabetes, hepatic disease or HIV).

Inclusion criteria patient sample. English speaking male and female volunteers classified as having severe brain injury and deemed able to make an informed decision to participate in the project. Severe brain injury was defined as satisfying at least one of the following criteria: (a) score of less than 9 on the Glasgow Coma Scale at time of injury (Teasdale & Jennett, 1974); (b) Post traumatic amnesia for at least 24 hours; (c) Loss of consciousness for 30 minutes or more following injury.

Exclusion criteria patient sample. Dysphasia.

Inclusion criteria control sample. English speaking male or female volunteers who know the patient and have given consent to participate. The aim was to match

participants on the following factors as far as possible: socio-economic status, education level, gender and age.

Exclusion criteria control sample. Same as the normative sample.

Recruitment procedures

The normative sample was recruited by sending information regarding the study to as many potential male and female participants as possible, via email, social networking sites and individual contact. For the social network site and emails, an initial invitation message or email was sent to a group of contacts and also individual contacts, giving information about the study. If they expressed interest in participating, they were sent the participant information sheet and asked to send an email to a specified email address.

The patient sample was recruited from two inpatient brain injury rehabilitation centres in the UK. The clinical team identified and referred participants into the project, and potential participants were given information about the study (by a member of the treating clinical team). Only patients deemed by the clinical team to have capacity to consent to this research project were invited to participate. The patient was given as long as s/he chose to decide (within the period of the study). If the patient indicated that they were interested in participating, the researcher met with the patient, and answered any questions about the project. If the patient was willing to participate, written consent was obtained and testing then proceeded.

To obtain the matched control group, patients were asked to nominate a friend or relative that may be willing to take part in the research, after they had completed the questionnaires. If deemed appropriate by the clinical team, the friend/family member was then contacted by the clinical team and invited to participate in the study. Relatives or friends who were interested in participating were given time to ask any questions they had about the project. Relatives and friends who consented to participate were then tested. Testing was carried out by sending questionnaires by post, completing it over the phone, or testing participants face-to-face. Participants were asked their preferred method of completion.

For the exploratory part of the analyses, the sample was established using data from patients in the first group comparison. The patients were grouped by whether the clinical team, including the psychiatrist, deemed them able, or not able, to make major welfare decisions (e.g. concerning need for care support, accommodation, relationships and work) in their daily lives. Those deemed not to have the ability to make important decisions were defined as requiring treatment under incapacity legislation (Adults with Incapacity Act Scotland, 2000, and the Mental Capacity Act, 2005). The criteria under which this decision was invoked included the following: the individual having an understanding of the decision to be made; memory for the relevant information; ability to weigh up options; communicate the decision; and act on the decision. The definitions of capacity are so similar between the acts as to allow the categorisation to be relatively accurate. The Consultants at each service used BIRT wide paperwork to decide on capacity and therefore were using the same criteria across centres.

Design

This study adopted a case-control and correlational design. It was split into three stages.

Phase 1. Existing CETs were evaluated by the project team (Main Researcher CT, Academic Supervisor JE, Field Supervisor BON). From these tests, questions that were deemed culture fair were included or modified to include in our new measure. New questions were also developed based on the following criteria:

1. Have no exact answer.
2. Rely on general knowledge that most people have (i.e. not a question some might get correct due to specific experience relevant to the question.).
3. Be fair to those in all cultures as far as possible.
4. Considered likely to produce a range of answers in the healthy population, but not so great a range as to render it unlikely to discriminate people with brain injury from controls.
5. Be relevant to everyday objects/activities.

In addition to the above criteria for questions, we also wanted a range and balance of questions, which included weight, size, quantity, time/duration and distance. To provide a check on whether the included items were culture fair, nine colleagues of the research team (two from India, and one from each of Qatar, Poland, Slovakia, Netherlands, Spain, Chile, and Pakistan) were invited to comment on whether they considered the items to be appropriate for their culture/context. Appropriate items were defined as ones that are understandable, to which it seems likely that an average person in their

culture/context could give a reasonable estimate, and for which the majority of people do not have highly specific knowledge. The most important feature of this test is that estimates are not based on detailed pre-existing knowledge of the specific focus of the question, but sufficient knowledge is available to make a reasonable estimate.

Phase 2. A normative data sample was collected via a number of methods including via an online survey tool, by sending questionnaires via post and testing participants face-to-face. An initial invitation email was sent to potential participants giving information about the study. If they wanted to participate, they were asked to send a reply email informing the researcher that they would like to take part. The researcher then sent an email back to the participant with a link to the online survey tool (Qualtrics). Participants were then taken to an introductory page, which explained what they were being asked to do. They then followed instructions to complete the demographic survey and CFACE. Freepost letters were sent to participants who may not access computers or asked for a paper version. Finally, the researcher also administered the questionnaire face-to-face, to as many people as possible who volunteered.

A sub-sample of 30 participants was invited to complete the test again in order to examine test-retest reliability. This was administered no more than one month after the first test. All participants were asked when they first completed the test whether they would be willing to be contacted again to complete the test on a further occasion and the sample of 30 were randomly selected from those who agreed.

CETs are generally scored in terms of amount of deviation from typical answers. Percentile scores for each item were generated based on the answers from the normative sample. For each participant, an error score was assigned for each answer based on percentile data and the error scores were totalled for the 30 items. Participants' demographic information including age, gender, years of education, occupation, handedness, and a measure of socioeconomic deprivation based on postcode were gathered. The relationship between scores and demographic data were assessed to see if any factors correlated with performance on the test.

Phase 3. The developed questions were administered to a group of people with severe brain injury and compared to a healthy, matched control group. In order to obtain a more closely matched control group, relatives and friends of participants were invited to take part in the study, where possible.

In addition to the CFACE, participants also completed the Dysexecutive Questionnaire (DEX; Burgess, Alderman, Evans, & Emslie, 1996), a subtest of the Behavioural Assessment of the Dysexecutive Syndrome (BADS), to assess concurrent validity of the newly developed measure (whether it correlates well with this already well-validated measure which demonstrates sensitivity to everyday problems experienced by those with brain injury). The DEX provides a more general measure of the impact of executive functioning difficulties, and includes items relevant to making decisions in everyday life. Therefore we might expect aspects of this measure (those relevant to functions used in estimation), in particular those from the independent rater, to correlate with our test.

CFACE performance was then assessed to explore whether it distinguished between those who professionals deemed as having capacity to make important decisions about their welfare and those who do not. This would constitute a real life, clinical issue regarding the importance of decision-making abilities in those with brain injury and would help determine the ecological validity of this test, and therefore whether this measure may be useful in a clinical setting.

Furthermore, ten items that were most sensitive to brain injury were selected based on differences between the brain injury group and both the matched control group and normative sample, on each question, to determine which items were most sensitive to patients with brain injury. These items could form a short version of the test, which could potentially lend itself to further assessment regarding its validity and reliability in different populations.

Measures

Participants with brain injury and matched controls.

- ***Culture Fair Assessment of Cognitive Estimation (CFACE).*** See Appendix 3.3.
  - This questionnaire was developed for this study and includes 30 culture fair cognitive estimation questions developed according to a specified set of criteria.
- ***Dysexecutive Questionnaire (DEX, Burgess et al., 1996)***
  - This questionnaire is a part of the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al., 1996), and includes 20 items

designed to sample emotional, motivational, behavioural and cognitive changes in someone with Dysexecutive Syndrome.

- The brain injury group completed the self-rater DEX questionnaire. A relative, friend or member of the care staff who knew the patient well completed the independent-rater DEX evaluating level of executive problems of the patient. A global measure of insight into post-injury deficits was obtained by subtracting a patient's self-ratings from those of an independent rater. In addition, the level of correlation between scores on the CET and the executive cognitive subscale (Simblett & Bateman, 2011) of the DEX was examined.

Data from the following measures, routinely available within the clinical service from which patients were recruited, was collected to characterise the patient sample:

- ***Behavioural Assessment of Dysexecutive Syndrome (BADS, Wilson et al., 1996)*** is comprised of six subsets and is designed to assess skills and demands of everyday life and is sensitive to capacities affected by frontal lobe damage.
- ***The Test of Premorbid Functioning (TOPF, Wechsler, 2009)*** is a test that enables clinicians to estimate an individual's level of intellectual functioning before the onset of the brain injury or illness.
- ***The Wechsler Adult Intelligence Test- fourth edition (WAIS-IV, Wechsler, 2008a)*** was designed to measure intelligence in adults and is often used to assess cognitive functioning after brain injury. It is comprised of 10 core subtests and 5 supplementary subtests.

Justification of Sample Size

Normative sample. Consistent with previous publications, we aimed to collect a normative sample size of at least 200. From the literature, it appears that four key factors influence performance on CETs: gender, education, intelligence/IQ and socio-economic status. Demographic information (gender, education and socio-economic status – estimated via postcode and occupation) was therefore collected and analysed to examine the impact of these factors on performance. We aimed to recruit a heterogeneous sample from the population representing a range of subgroups within each of these factors.

Patient sample. The majority of the previous literature has tested construct validity, looking at how well the CET distinguishes between those with and without brain injury. A selection of these was evaluated. The papers assessed patients with a range of neurological issues including neurological, neurosurgical and neuropsychiatric problems. The effect sizes calculated using Cohen's d, ranged from 0.05-2.90 with an average effect size of 1.0. For the first comparison therefore, the calculation was powered on the basis of an effect size of 0.8, as this is deemed reasonable given previous publications. Our alpha level was 0.05 (one-tailed), and using G-power to calculate sample size, we aimed for minimum of 21 participants in each group.

For the second exploratory hypothesis, examining whether there is a significant difference between those with brain injury who are deemed to have capacity to make decisions, and those who are not, it was not clear what the effect size would be as no previous studies have examined this; however, it would need to be large to be clinically

useful. Therefore, this part of the study was also powered on the basis of a large effect size.

Ethical Approval

Ethical approval for this study was obtained via the University of Glasgow College of Medical Veterinary and Life Sciences ethics panel (Reference number 200140002; Approval date: 14th November 2014). Sponsorship and approval for the project was also provided from the Brain Injury Rehabilitation Trust. See Appendix 3.1.

Results

Two hundred and thirty seven volunteers comprised the normative sample. 172 people altogether completed the questionnaire online and 65 people completed the questionnaire on paper or face-to-face. Table 1 shows the sample characteristics.

The Scottish Index of Multiple Deprivation (SIMD) postcode excel lookup was used as a tool to derive the deprivation status of the area in which participants lived. This converted postcode data into ranks from 1 (most deprived) to 6505 (least deprived), and also into quintiles (1-5) and deciles (1-10). For the purpose of this evaluation, the quintile postcode data were used. This data relates to socio-economic status (SES), a factor previously shown to influence performance on the CET. SIMD could only be applied to participants living in Scotland, which comprised 69.6% of the normative sample, and 67.4% of the patients and matched controls. Using this sub-sample was helpful in determining influence of SES on CFACE performance. Occupation was coded according to 5 categories in the National Statistics Socio-economic Classification Volume 3 (NS-SEC 2010, see Appendix 3.7).

Table 1 Normative sample characteristics

Sample Characteristics (N=237)	N (%)	M (sd)	Range
Age	237	37.0 (12.9)	18-66
Gender	237 M 94 (39.7) F 143 (60.3)		
Age left education *(categories)*	233 (98.3)		
16 or before	10 (4.3)		
17-19	21 (9.0)		
After 18	202 (86.7)		
Still studying	44 (18.6)		
Mean age left education	233 (98.7)	24.1 (5.0)	14-41
Level of education	237 (100)		
Not specified/none	2 (0.8)		
GCSE/Standard grade/Olevel	10 (4.2)		
Higher/Alevel	14 (5.9)		
Post-school qualification	21 (8.9)		
Undergraduate	105 (44.3)		
Postgraduate	85 (35.9)		
Occupation	237		
1. *Higher managerial, administrative and professional occupations*	171 (72.5)		
2. *Intermediate occupations*	26 (11.0)		
3. *Small employers and own account workers*	4 (1.7)		
4. *Lower supervisory and technical occupations*	4 (1.7)		
5. *Semi-routine and routine occupations*	14 (5.9)		
6. *Never worked, long-term unemployed*	5 (2.1)		
7. *Student*	9 (3.8)		
8. *Retired not specified*	3 (1.3)		
Handedness	237 (100)		
Right-handed	205 (86.5)		
Left-handed	28 (11.8)		
Ambidextrous	4 (1.7)		
Index of Multiple Deprivation (based on Scottish postcodes, 1 most deprived to 5 least deprived).	165 (69.6)		
Quintile			
1	22 (13.3)		
2	19 (11.5)		
3	20 (12.1)		
4	37 (22.4)		
5	67 (40.6)		

*72 (30.4%) were non-Scottish postcodes, therefore non-interpretable according to SIMD.

CFACE Scoring Method

Answers to questions were given using a range of metrics and measurements, and so for each item, answers were converted to the metric considered most relevant to the question. The means, standard deviations and medians for each of the 30 items were then recorded. See Table 2 for item characteristics.

Table 2 Normative sample characteristics (N= 237)

CFACE ITEM	Mean (sd)	Median	Range	Skewness (SE)	Normally distributed?
1*	159.4 (99.8)	150.0	10-1000	3.4 (.16)	N
2*	7.2 (1.7)	7.0	3-12	-.14 (.16)	Y
3 (cm)	210.8 (135.1)	200.0	30.5-1000	3.4 (.16)	N
4*	23.9 (12.9)	20.0	4-100	2.3 (.16)	N
5 (g)	222.2 (990.5)	56.7	0.05-10000	8.7 (.16)	N
6 (m)	62.06 (217.5)	20.0	2-2000	6.8 (.16)	N
7*	1487.0 (6709.8)	600.0	8-100000	13.6 (.16)	N
8 (g)	1614.6 (1816.8)	1200.0	12-20000	6.4 (.16)	N
9 (s)	280.5 (236.2)	240.0	10-1800	3.3 (.16)	N
10*	64.8 (108.7)	40.0	0-1000	6.4 (.16)	N
11 (mph)	35.5 (16.0)	35.0	3.1-110	1.3 (.16)	Y
12 (km)	30.0 (14.1)	29.0	0.1-128.7	1.7 (.16)	N
13*	39.5 (31.3)	32.0	5-300	3.8 (.16)	N
14 (l)	3.6 (2.9)	3.0	0.06-30	4.5 (.16)	N
15 (feet)	165.1 (5.2)	165.1	130-195.6	-.9 (.16)	Y
16 (g)	1293.2 (1774.9)	907.2	7-20000	6.4 (.16)	N
17*	5496.5 (65218.5)	300.0	20-1000000	15.1 (.16)	N
18 (s)	105.0 (83.9)	65.0	3-600	2.1 (.16)	N
19 (ml)	196.8 (182.1)	150.0	5-1500	2.6 (.16)	N
20 (cm)	96.0 (76.0)	91.4	38.1-800	7.4 (.16)	N
21 (g)	381.1 (490.2)	226.8	0.05-4535.9	4.2 (.16)	N
22 (min)	94.1 (79.7)	80.0	20-1140	9.9 (.16)	N
23 (km)	9.4 (64.7)	4.8	0.03-1000	15.3 (.16)	N
24*	88.5 (148.4)	60.0	10-2000	10.2 (.16)	N
25 (g)	11706.5 (58983.1)	5000.0	15-900000	14.4 (.16)	N
26*	10621.7 (130111.5)	600.0	18-2000000	15.3 (.16)	N
27 (mph)	39.7 (26.2)	35.0	1-200	14.8 (.16)	N
28 (min)	14.7 (17.0)	10.0	1-180	2.3 (.16)	N
29 (kg)	7672.4 (25108.7)	2500.0	0-254000	6.8 (.16)	N
30*	447.7 (1028.3)	130.0	0-6000	3.9 (.16)	N

*Quantity question e.g. *how many*

A normative distribution curve was derived for the range of answers given for each estimate question. A scoring method was derived using percentiles due to the majority of questions (n=27) not showing normal distribution. The percentiles for each item's actual responses were examined and error scores assigned to the following ranges of percentiles. Responses from 26th to 75th percentile were considered within the normal range and were awarded 0 points. Responses that ranged from the 21st to 25th or 76th to 80th percentile were awarded 1 point. Scores within the ranges of 16th to 20th or 81st to 85th percentiles were awarded 2 points. Those scores in the 11th to 15th percentile range or 86th to 90th range scored 3 points. Scores within the 91st to 95th or falling within the 5th to 10th percentile range were considered extreme and scored 4 points. Finally, responses less than the 5th percentile or more than the 95th percentile were considered very extreme and awarded 5 points. Therefore the higher the score, the further away the answer was from the norm. See Appendix 3.6 for error score conversion tables.

The error scores were applied to actual scores for each item in the normative sample and added together to assess overall deviation. This total error score (TES) was derived for each participant. The higher the total score, the bigger the distance of that individual's performance from the average response. For missing data (1.4% of the total number of answers), mean values were inputted in order that total scores for every participant could be used. One extreme outlier was identified (TES= 122) by examining the range of scores and box plots, and removed from the dataset. For the 236 participants in the normative sample, the mean TES was 47.6 with a standard deviation of 15.4, the median value was 47.0 and the range of TES was 15-95.

Spearman's rank order correlations were conducted to examine the correlations between TES and a range of demographic variables to determine whether any factors correlated with performance on the test. There were significant negative correlations between TES and deprivation quintile derived from postcode data (rho= - .171, p= .009), meaning that those living in more deprived areas were more likely to have higher TES. A significant negative correlation was also found between TES and level of education, so the higher the level of education, the lower the TES (rho= - .162, p= .013). A Mann-Whitney U test was conducted for gender group, showing that females (Mdn= 48.5) had significantly higher TES than males (Mdn= 46.0; U= 5495.5, Z=-2.296, p= .022, r= 0.16). A Kruskal-Wallis H test showed no significant differences in TES between handedness groups (χ^2 (2)= 2.37 p= .306). No significant correlations were found between TES and age (rho=- .116, p= .075), occupation (rho= .017, p= .794), or age left education (rho=- .109, p= .099). The CFACE was completed in different formats, and a comparison of groups showed no differences in TES whether the questionnaire was completed online or on paper/face-to-face (U= 5236, Z= -.752, p= .452).

Group Comparisons

Twenty-three patients with moderate to severe brain injury were recruited. Sixteen control participants were recruited from amongst friends or family members of the patients. They were deemed more likely to match patient characteristics in terms of SES, knowledge and education level; factors previously associated with performance on this test. For seven of the BI patient group, a friend/relative was not available and therefore for each patient, a control participant from the normative sample was selected with the aim of matching participants as closely as possible on as many demographic variables as possible, paying particular attention to factors influencing performance in

this current study. Therefore factors were prioritised in the following in order: level of education, postcode, gender, age left education, occupation and age. Controls were therefore selected by taking each factor above, in order, and systematically moving down the list of participants until all factors were matched for each of the BI participants. This process was completed without reference to the TES. Table 3 provides a summary of the BI group and control group characteristics.

Table 3 Characteristics of BI patient group and control group

BI group (N= 23)	**N (%)**	**M (sd)** ***Range***	**Control group (N=23)**	**N (%)**	**M(sd)** ***Range***
Age	23	41.4 (12.0) *18-63*	**Age**	23	46.0 (13.7) *18-70*
Gender	M 19 (82.6) F 4 (17.4)		**Gender**	M 9 (39.1) F 14 (60.9)	
Age left education			**Age left education**		
16 or before	13 (56.5)		*16 or before*	12 (52.2)	
17-19	5 (21.7)		*17-19*	6 (26.1)	
After 18	3 (13.0)		*After 18*	5 (21.7)	
Still studying	2 (8.7)		*Still studying*	0	
Mean age left education	23 (100)	17.4 (3.0) *14-28*	*Mean age left education*	23 (100)	17.8 (4.1) *14-33*
Level of education			**Level of education**		
Not specified/none	5 (21.7)		*Not specified/none*	6 (26.1)	
GCSE/Standard grade/Olevel	10 (43.5)		*GCSE/Standard grade/Olevel*	5 (21.7)	
Higher/Alevel	3 (13.0)		*Higher/Alevel*	4 (17.4)	
Post-school qualification	3 (13.0)		*Post-school qualification*	4 (17.4)	
Undergraduate	0		*Undergraduate*	4 (17.4)	
Postgraduate	2 (8.7)		*Postgraduate*	0	
Occupation			**Occupation**		
1. Higher managerial, administrative and professional occupations	0		*1. Higher managerial, administrative and professional occupations*	5 (21.7)	
2. Intermediate occupations	2 (8.7)		*2. Intermediate occupations*	4 (17.4)	
3. Small employers and own account workers	4 (17.4)		*3. Small employers and own account workers*	1 (4.3)	
4. Lower supervisory and technical occupations	3 (13.0)		*4. Lower supervisory and technical occupations*	0	
5. Semi-routine and routine occupations	11 (47.8)		*5. Semi-routine and routine occupations*	9 (39.1)	
6. Never worked, long-term unemployed	2 (8.7)		*6. Never worked, long-term unemployed*	4 (17.4)	
7. Student	1 (4.3)		*7. Student*	0	
8. Retired not specified	0		*8. Retired not specified*	0	
Handedness			**Handedness**		
Right-handed	18 (78.3)		*Right-handed*	20 (87.0)	
Left-handed	4 (17.4)		*Left-handed*	3 (13.0)	
Ambidextrous	1 (4.3)		*Ambidextrous*	0	
Index of Multiple Deprivation (based on Scottish postcodes)	15 (65.2)		**Index of Multiple Deprivation** (based on Scottish postcodes)	16 (69.6)	
Quintile			*Quintile*		
1	7 (46.7)		*1*	5 (31.25)	
2	5 (33.3)		*2*	4 (25.0)	
3	1 (6.7)		*3*	2 (12.5)	
4	1 (6.7)		*4*	2 (12.5)	
5	1 (6.7)		*5*	3 (18.75)	
BADS (overall age corrected standardised score)	*14 (60.9)	69.1 (17.2) *38-98*			
TOPF (estimate of FSIQ)	20 (87.0)	95.3 (15.4) *66.2-128*			
WAIS verbal comprehension	22 (95.7)	82.4 (17.5) *54-132*			
WAIS perceptual reasoning	21 (91.3)	82.0 (11.5) *60-100*			
WAIS working memory	22 (95.7)	83.7 (14.1) *63-122*			
WAIS processing speed	20 (87.0)	69.3 (13.0) *50-94*			
WAIS FSIQ	20 (87.0)	75.5 (13.3) *53-106*			

*BADS data only available for patients recruited from one of the rehabilitation centres from where participants were recruited.

BI vs. Matched Controls. The distributions of the TES for both the BI and control group were examined by observation of histograms, boxpots, skewness, kurtosis scores and Shapiro-Wilk. Results showed that scores were approximately normally distributed for the BI sample but not- normally distributed for the control sample. Therefore non-parametric analyses were used.

Comparison of means using a Mann-Whitney U test, were conducted to explore whether there were any significant differences between groups on the factors associated with performance on the test (i.e. gender, level of education and postcode). There were no significant differences between groups on level of education (U= 236.0, Z=-. 644, p= .519) or postcode data (U= 222, Z= - .963, p= .335). However a Chi-Square test showed a significant association between gender and group (X (1)= 9.127, p= .003), with more males in the BI group (n=19) and more females (n=14) in the control group.

Table 4 TES for BI and control groups

Total error scores	*BI*	*Control*
n	23	23
Mean (sd)	77.56 (19.92)	54.48 (14.52)
Median	76.00	53.00
Range	50-129	36-97

A Mann-Whitney U test was conducted to assess whether there were any differences in scores between the BI and control group without consideration of gender. A significant difference between groups was found with the BI group (Mdn= 76.0) demonstrating a higher TES than the control group (Mdn= 53.0; U = 82.5, Z= -4.00, p= .001 r= 0.59).

In a further exploration, Mann-Whitney U tests were conducted and showed that there were no significant differences between males and females in TES within the BI group

(U = 26.0, Z=- .974, p= .330, r= 0.2) or control group (U = 32.5, Z=-1.926, p= .054, r= 0.4). A comparison was also conducted between males and females in the whole group (BI and controls) that showed no significant differences in TES (U=238.0, Z= -.315, p= .752, r= 0.05). Given that there were no statistical differences in performance between males and females within each group, as well as no effect of gender evident in the whole group (where there were more female controls and more male patients), it appeared that differences between the groups could not be explained by gender..

Exploratory comparison. The BI patient sample was categorised into groups based on their ability to make important welfare decisions. Table 5 shows group characteristics.

Table 5 Characteristics of the BI groups according to their ability to make welfare decisions

Deemed able to make welfare decisions (N= 11)	**N (%)**	**Mean (sd)** ***Range***	**Deemed NOT able to make welfare decisions** (N=12)	**N (%)**	**Mean (sd)** ***Range***
Age		41.8 (12.9) *20-63*	**Age**		41.0 (11.6) *18-59*
Gender	M 9 (81.8) F 2 (18.2)		**Gender**	M 10 (83.3) F 2 (16.7)	
Age left education			**Age left education**		
16 or before	7 (63.6)		*16 or before*	8 (66.7)	
17-19	2 (18.2)		*17-19*	3 (25.0)	
After 18	2 (18.2)		*After 18*	1 (8.3)	
Level of education			**Level of education**		
Not specified/none	1 (9.1)		*Not specified/none*	4 (33.3)	
GCSE/Standard grade/Olevel	7 (63.6)		*GCSE/Standard grade/Olevel*	3 (25.0)	
Higher/Alevel	0		*Higher/Alevel*	3 (25.0)	
Post-school qualification	2 (18.2)		*Post-school qualification*	1 (8.3)	
Undergraduate	0		*Undergraduate*	0	
Postgraduate	1 (9.1)		*Postgraduate*	1 (8.3)	
Occupation			**Occupation**		
1. Higher managerial, administrative and professional occupations	0		*1. Higher managerial, administrative and professional occupations*	0	
2. Intermediate occupations	1 (9.1)		*2. Intermediate occupations*	1 (8.3)	
3. Small employers and own account workers	2 (18.2)		*3. Small employers and own account workers*	2 (16.7)	
4. Lower supervisory and technical occupations	3 (27.3)		*4. Lower supervisory and technical occupations*	0	
5. Semi-routine and routine occupations	4 (36.4)		*5. Semi-routine and routine occupations*	7 (58.3)	
6. Never worked, long-term unemployed	1 (9.1)		*6. Never worked, long-term unemployed*	2 (16.7)	
7. Student	0		*7. Student*	0	
8. Retired not specified	0		*8. Retired not specified*	0	
Handedness			**Handedness**		
Right-handed	10 (90.9)		*Right-handed*	8 (66.7)	
Left-handed	1 (9.1)		*Left-handed*	3 (25.0)	
Ambidextrous	0		*Ambidextrous*	1 (8.3)	
Index of Multiple Deprivation (based on Scottish postcodes)	N=6		**Index of Multiple Deprivation** (based on Scottish postcodes)	N=9	
Quintile			*Quintile*		
1	3 (50.0)		*1*	4 (44.4)	
2	1 (16.7)		*2*	4 (44.4)	
3	1 (16.7)		*3*	0	
4	0		*4*	1 (11.1)	
5	1 (16.7)		*5*	0	
BADS (overall age corrected standardised score)	*6 (54.5)	68.7 (21.8) 38-98	**BADS (overall age corrected standardised score)**	8 (66.7)	69.4 (14.5) 53-86
TOPF (estimate of FSIQ)	10 (90.9)	97.3 (11.6) 87.5-126.9	**TOPF (estimate of FSIQ)**	10 (83.3)	93.3 (18.8) 66.2-128
WAIS verbal comprehension	11 (100.0)	80.2 (10.1) 68-98	**WAIS verbal comprehension**	11 (91.7)	84.6 (23.1) 54-132
WAIS perceptual reasoning	10 (90.9)	82.0 (13.9) 60-100	**WAIS perceptual reasoning**	11 (91.7)	82.0 (9.6) 65-98
WAIS working memory	11 (100.0)	83.0 (10.4) 71-100	**WAIS working memory**	11 (91.7)	84.4 (17.5) 63-122
WAIS processing speed	9 (81.8)	75.9 (10.7) 56-94	**WAIS processing speed**	11 (91.7)	63.8 (12.6) 50-86
WAIS FSIQ	9 (81.8)	75.2 (12.7) 56-94	**WAIS FSIQ**	11 (91.7)	75.7 (14.5) 53-106

*BADS data only available for patients recruited from one of the rehabilitation centres from where participants were recruited.

Those deemed able to make these decisions (n=11) were compared with those not deemed able to make these decisions (n=12). There were no significant differences between the groups in gender (U = 65.0, Z=-. 094, p= .925), level of education (U = 59.0, Z=- .45, p= .650) or quintile postcode data (U = 54.5, Z= - .74, p= .461). A Mann-Whitney U test was then carried out to determine whether there were any differences between capacity groups on CFACE performance. Results showed that there were no significant differences between the two groups, with those deemed to have capacity scoring a median TES of 80.0 (mean =76.82, sd= 15.45), and those deemed not to have this capacity scoring a median TES of 72.5 (mean= 78.25, sd = 24.0; U= 63.5, p= .878, r= 0.03).

Differences between groups on other clinical measures were also explored. There were no significant differences between groups on the BADS overall age corrected standardised score (U = 23.0, Z=-.130, p= .897, r=0.03), TOPF estimate of FSIQ (U = 44.5, Z= - .416, p= .677, r= 0.09), WAIS verbal comprehension (U =58.0, Z= - .164, p= .869, r= 0.03), WAIS perceptual reasoning (U = 52.5, Z= - .177, p= .860, r= 0.04), WAIS working memory (U = 59.5, Z= - .066, p= .948, r=0.01), or WAIS FSIQ (U = 49.0, Z= - .038, p= .970, r=0.008). There was however, a significant difference between groups on WAIS processing speed (U = 22, Z= -2.10, p= .036, r= 0.44) with the group deemed to have capacity to make welfare decisions (Mdn= 76.0, mean= 75.89, sd= 10.69) performing better overall than those without this capacity to make welfare decisions (Mdn= 68.0, mean= 63.82, sd= 12.58).

When comparing these groups on the DEX ratings, there were no differences with the DEX-self rating (U = 53.0, Z= - .802, p= .423, r= 0.17), however there was a significant

difference between groups on the DEX-independent rating (U = 31.5, Z= - 2.13, p= .033, r= 0.44) with those deemed to be able to make welfare decisions (Mdn= 27.0, mean=29.64, sd=12.81) receiving a lower score on the DEX than those not deemed able to make welfare decisions (Mdn= 40.5, mean=42.50, sd= 13.85). A global measure of insight into post-injury deficits was obtained by subtracting a patient's self-ratings from those of an independent rater. There were no significant differences between groups on these scores (U= 64.0, Z= - .125, p= .928).

Correlational Analyses

The DEX was administered at the same time as the CFACE (to ensure time of testing did not interfere with results). Correlations between scores on the CFACE and the DEX were explored. Spearman's rank order correlations for ordinal data were conducted. There was a medium positive correlation between the self and independent ratings on the DEX, however this did not reach significance (rho= .392, p= .065). Similarly, there were no significant correlations between the DEX self-rating and the patient TES (rho= .187, p= .393), the DEX independent- rating and the patient TES (rho= .002, p= .993), or between difference scores on the DEX self and DEX independent rating and the patient TES (rho= .093, p= .673).

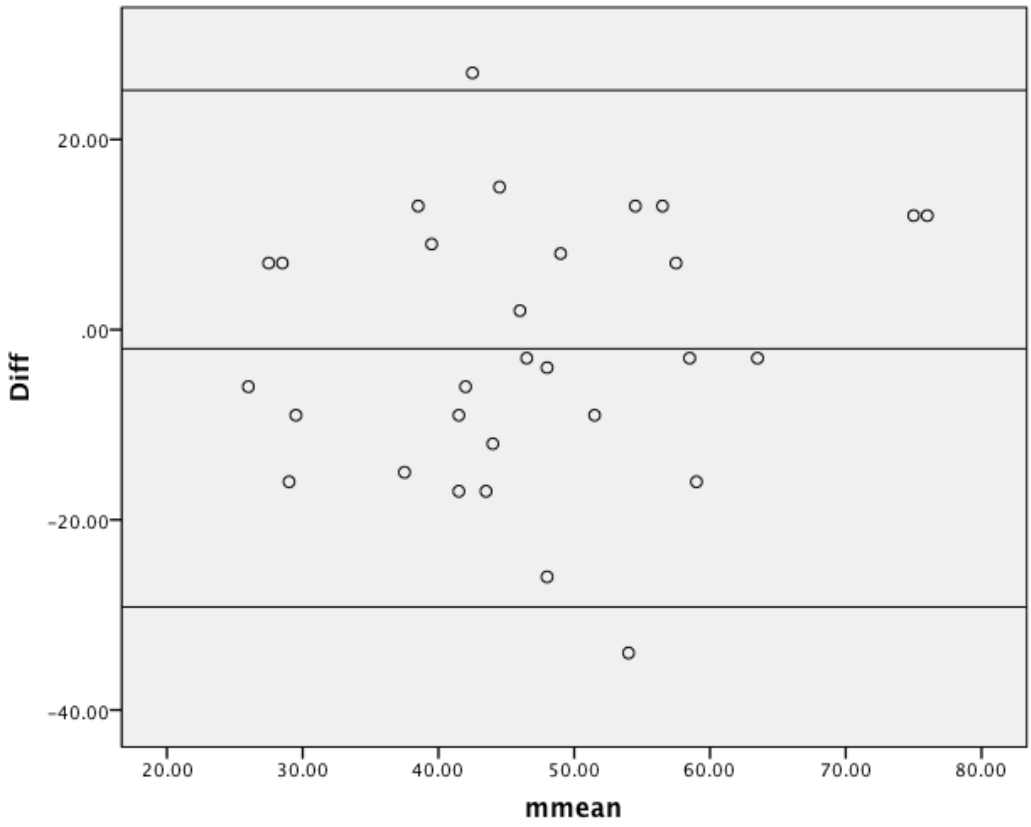
Correlations with the DEX executive-cognition component. Concurrent validity against executive-cognition components of the DEX, identified by Simblett and Bateman (2011) using principal component analysis, was assessed. Ratings for items 1, 2, 6 and 18 were added to give a total executive-cognition score. Spearman's rank order correlations were then conducted to assess any relationships between both the self and

independent rating total executive-cognition scores with the patient TES. No significant relationship was found between the self-rating on executive-cognition measures and the patient TES. For the independent rating score, there was a medium-large positive correlation though this did not reach significance (rho= .404, p= .056).

Test-re-test reliability. The test-retest correlation between the two time points was examined, with an expectation that the correlation should be at least 0.7. Thirty participants from the normative sample completed the CFACE on a second occasion, between two weeks and a month after having first completed it. TES were calculated for each participant at time 1 and time 2. Spearman's rank order correlations were conducted to explore correlations between these two time points. The two time points were significantly correlated (rho= .490, p= .006). A Wilcoxon signed ranks test was then conducted to determine any differences between scores at the two time points. There was no significant difference between the mean scores at the different time points (Z= .710, p= .478).

The Bland and Altman (1986) statistical method of assessing agreement between clinical assessment methods was used to determine 95% confidence limits for the difference between the two measurement points. A Bland-Altman plot was constructed (see Figure 1). A one-sample t-test was conducted to explore systematic deviation from the mean (0 difference). The t-score was not statistically significant (t= .807, p= .427), therefore indicating no systematic bias in scores across the mean. However, the level of agreement between time 1 and time 2 can be seen to be poor given the range of difference scores and likely indicates a low level of reliability between time points.

Figure 1 Bland- Altman Plot showing mean difference in scores and upper and lower confidence limits



The Cronbach's alpha for the 30 items on the CET was .70, which is considered to be an acceptable level of internal reliability.

Selection of Sub-Items Most Sensitive to BI

Mann-Whitney U tests were conducted to explore differences between the BI group and the matched control group, and the BI group and the normative sample on each question, to determine which items were most sensitive to the patients with brain injury. Table 6 shows significance values.

Table 6 Significance values for each CFACE item

CFACE item	BI vs. control group 1 (sig.)	Effect size (r)	BI vs. normative sample	Effect size (r)
1	*.013**	*0.37*	*.000***	*0.23*
2	*.066*	*0.27*	*.002***	*0.19*
3	*.415*	*0.12*	*.040**	*0.13*
4	*.193*	*0.19*	*.040**	*0.13*
5	*.018**	*0.35*	*.001***	*0.22*
6	*.726*	*0.05*	*.014**	*0.15*
7	*.001***	*0.62*	*.000***	*0.29*
8	*.050**	*0.29*	*.010***	*0.16*
9	*.312*	*0.15*	*.031**	*0.13*
10	*.220*	*0.18*	*.175*	*0.08*
11	*.131*	*0.22*	*.003***	*0.19*
12	*.304*	*0.15*	*.000***	*0.23*
13	*.721*	*0.05*	*.122*	*0.10*
14	*.460*	*0.11*	*.844*	*0.01*
15	*.055*	*0.28*	*.000***	*0.25*
16	*.390*	*0.13*	*.043**	*0.13*
17	*.732*	*0.05*	*.259*	*0.07*
18	*.268*	*0.16*	*.016**	*0.15*
19	*.319*	*0.15*	*.249*	*0.07*
20	*.211*	*0.18*	*.459*	*0.05*
21	*.039**	*0.30*	*.026**	*0.14*
22	*.369*	*0.13*	*.015**	*0.15*
23	*.027**	*0.33*	*.065*	*0.11*
24	*.148*	*0.21*	*.038**	*0.13*
25	*.559*	*0.09*	*.001***	*0.20*
26	*.735*	*0.05*	*.122*	*0.10*
27	*.167*	*0.20*	*.031**	*0.13*
28	*.009***	*0.39*	*.001***	*0.20*
29	*.406*	*0.12*	*.009***	*0.16*
30	*.061*	*0.28*	*.006***	*0.17*

**indicates significance at p< .05*
***indicates significance at p< .01*

10 items were selected as those more sensitive to BI based on the p-values, Z-scores and effect sizes across the 2 comparisons. They included the following (Table 7):

Table 7 Items most sensitive to BI

CFACE item	Question	Dimension
1	*How many steps does it take for an adult to walk 100m (11yards/ 300 feet)?*	*Quantity*
2	*How many TV programs are there, on average, on any one channel between 6 and 11pm?*	*Quantity*
5	*How much does a hen's egg weigh?*	*Weight*
7	*How many times on average do we blink in an hour?*	*Quantity*
8	*How much do 12 average sized bananas weigh?*	*Weight*
11	*How fast do lions run?*	*Speed*
15	*How tall is the average woman?*	*Length*
21	*How heavy is a man's cotton shirt?*	*Weight?*
28	*How long would it take an adult to write a one-page letter?*	*Time*
30	*How many times can you wash your hands from a bar of soap?*	*Quantity*

A Mann-Whitney U test was also conducted to explore differences between the two BI groups on each question. Only one question (qu. 21; How heavy is a man's cotton shirt?) showed a significant difference between groups.

Discussion

Summary of Findings

This research was seeking to develop a new culture fair assessment and investigate its validity and reliability. Questions were developed taking into account feedback on cultural relevance from nine colleagues from eight different countries. A scoring method was designed based on a normative sample of healthy participants and applied to the scores given by a brain injury and control group. The results confirmed previous evidence that those with brain injury perform significantly worse than those without brain injury on this test; however no significant differences were found between two brain injury groups categorised according to whether clinicians deemed them able to make important welfare decisions. The effect size in this comparison was very small (r= 0.04) and therefore suggests that the CFACE does not distinguish between these groups. The inclusion of this exploratory comparison was a significant strength of this study in comparison to previous studies. Performance on the test was associated with gender, level of education and social deprivation status (measured according to postcode but not occupation). These findings support earlier studies showing higher education levels being associated with better performance on the CET (e.g. Axelrod & Millis, 1994, O'Carroll, Egan & MacKenzie, 1994). It also replicates previous literature that has not found correlations between CET performance and age (e.g. O'Carroll et al., 1994, Axelrod & Millis, 1994, Della Sala et al., 2003). No significant correlations were found between performance on the test and the DEX, an everyday measure of executive functioning, which suggests poor concurrent validity. The executive-cognition components of the DEX and performance on the CFACE showed a medium-large effect size (rho=0.404) albeit not reaching statistical significance. Test re-test reliability

showed a relatively weak positive correlation (.404) between the two time points, demonstrating relatively poor reliability or consistency of the measure across two time points. By exploring specific questions that were more sensitive to brain injury, 10 culture fair items were selected, which may be helpful in developing future tests of this kind.

Observations of Performance on the Test

One important observation was that many different measurement types (e.g. imperial and metric) were used. Often, it appeared that participants were unfamiliar with weight, distance or speed measurements and it is likely that this contributed to larger deviation scores in some cases. Therefore it is possible that some higher deviation scores resulted from lack of knowledge on measurements rather than a lack of appropriate reasoning or estimation abilities. The distinction between imperial versus metric item responses could potentially be seen to limit the extent to which the test is culture fair. However items were converted to the same measure after test completion, and therefore participants were not restricted in how they answered the questions. For two of the participants in the brain injury sample, the same number was repeated on a few different occasions for different items. This was not observed in the control sample.

The variation in responding based on gender was also an interesting observation and has been demonstrated previously by Della Sala et al. (2003). Like the current study, Della Sala et al. (2003) also found that men performed better than women overall. They commented that cognitive estimation tasks require an individual to be able to retrieve a congruent set of items from the archives of General Knowledge of the World (GKW). Women proved to be poorer than men on a wide range of GKW items in a study by Mariani, Sacco, Spinnler & Venneri (2002), which may also have been a reason for

their poorer CET performance. It is also possible that the gender difference in response could be an artifact of the fewer males in the normative sample (N= 94) than females (N= 143).

Previous Studies

Previous literature on CETs has largely concluded that the CET distinguishes those with brain injury versus healthy controls, particularly those with frontal lesions. MacPherson, Wagner, Murphy, Bozzali, and Cipolotti et al. (2014) devised two nine-item parallel versions of the CET and administered these to patients with frontal lobe lesions and healthy controls. They found that the frontal patients' error scores were significantly higher than the healthy control group on the two tests. Findings have been extended to a number of neurological deficits, which often demonstrate dysexecutive or frontal presentations. Brand, Kalbe, Fujiwara, Huber and Markowitsch (2003) assessed Alzheimer's patients, patients with Korsakoff's syndrome and healthy controls on a 16 item German version of the CET and found that both patient groups were strongly impaired compared to controls, scoring overall higher error scores. Roth, Pixley, Kruck and Garlinghouse (2012) explored performance on a 10 item version of the CET (developed by Axelrod & Millis, 1994) in patients with schizophrenia and found that this patient group performed more poorly than the healthy comparison group matched for age, gender and parental education. The findings in the present study largely replicate previous literature. There have been contradictory findings however, for example, Taylor and O'Carroll (1995) did not find evidence of poorer performance on the CET developed by Shallice and Evans (1979) in those with brain injury, with the exception of those with Korsakoff's syndrome. In this study, 370 patients with a range of neurological and psychiatric conditions such as head injury, brain tumour, ruptured

aneurysm, multiple sclerosis, dementia, encephalitis, Korsakoff's syndrome, anxiety and depression were compared with scores from 150 healthy controls. Different neurological presentations were explored separately. There were a range of severities, and for those with head injury for example, 74 were categorised as severe and 20 as moderate according to the duration of post-traumatic amnesia and also the initial Glasgow Coma Scale score.

The validity of the CET and its usefulness, however, cannot be determined based on a comparison between a patient and healthy control group. There are a number of different forms of validity, but in essence, *"Validity refers to the degree to which evidence supports the interpretation of test scores for their intended purpose; therefore the examination of a test's validity requires an evaluative judgement by the test user"* *WMS*-IV Technical and Interpretive Manual (2009). To be a clinically useful measure in this context, this measure should be able to relate functional everyday decision-making, which impacts on an individual's well being. As an exploratory part of this study, therefore, the patients with brain injury were categorised into two groups: one group deemed able to make important welfare decisions for themselves and one group who were not. However, there were no significant differences in performance between these two groups. No previous studies have assessed this.

A number of previous studies have also examined the relationship between performance on a CET with other valid neuropsychological tests, including measures of everyday executive functioning. Burgess, Alderman, Evans, Emslie and Wilson (1998) examined the relationship between the original CET (Shallice and Evans, 1978) and the Dysexecutive Questionnaire (DEX). They found no significant correlations between scores on the DEX and the CET. Similarly, Silverberg, Hanks and McKay (2007) conducted a study exploring the construct and ecological validity of the Biber cognitive

estimation test (BCET, Bullard et al., 2004). The BCET did not predict current functional status, measured using the Disability Rating Scale. Results from the present study are broadly consistent with these previous findings, though a subscale of the DEX did show a moderate (albeit non-significant) correlation with the CFACE, perhaps showing that the CFACE includes more specific cognitive-executive components of everyday life.

Study Strengths and Limitations

The developed questions were carefully considered in accordance to what they were intended to measure. The fact they were assessed regarding their culture fairness was a helpful addition to the design of the test. Another strength was the consideration of the clinical validity of the test in terms of its ability to demonstrate differences between patients with different levels of capacity to make decisions, and also its relationship with a validated test of everyday executive functioning. Studies only comparing performance of a brain injury group with a healthy control group on the CET are helpful as a starting point in developing a measure, however, they do not offer any information on the usefulness of the CET in clinical practice, nor provide information on the specificity of the CET to any particular brain or cognitive dysfunction. Therefore, it is important that future studies continue to examine more useful comparison samples for the intended purpose of this test.

This research was seeking to develop a new culture fair assessment, and the extent to which this measure meets this criterion needs to be evaluated in future research. This study did not provide details about the nationality, ethnicity or cultural identity of the participants; therefore there is no information on whether the groups recruited were

ethnically/culturally heterogeneous. It was observed however, that the groups mainly consisted of British Caucasian participants. Overall, this study did not reasonably evaluate the culture fair nature of the test in a culturally diverse sample and therefore, future research will be required to evaluate this. Regarding our criteria for new items being 'fair to those in all cultures as much as possible', it was discussed when developing the questions that there would be limitations to how 'culture-fair' any question was, given the complex nature of culture and also reflection on the culture of testing. We aimed to develop questions that the majority of people, who were able to access medical care, could provide a reasonable answer to. Although there were knowledge limitations within the project team on worldwide cultural issues, we attempted to broaden our understanding by asking others living around the world their views on the questions. Information however only came from a limited number of countries and individuals.

A useful addition to this study would have been to include the original CET (Shallice & Evans, 1978) as a way of exploring the convergent validity of the CFACE. It would have been helpful to assess this new measure against an already validated measure and is something that could be explored in the future.

Another limitation is the possibility that the online questionnaire resulted in individuals not reading the instructions clearly. Some participants left out questions, or answered 'depends' for a number of the questions. This may have resulted in a number of the questionnaires being completed without much consideration or reasoned thought. It may also suggest that this scale does not work as an online questionnaire. Questionnaires were completed in different formats, however a comparison of groups showed no differences in TES whether completed online or on paper/face-to-face. A further limitation was that as recruitment was open to participants outside of Scotland this

meant that a proxy measure of socio-economic status based on postcode was only available for a proportion of the sample.

Although a reasonable normative sample size was gathered, it was not equally spread with regards to levels of education and SES, showing higher overall education levels. This meant that scoring was developed based on a sample that may be less generalisable to other populations. To ensure that the normative sample was representative, it may have been helpful to spend time visiting different areas of Scotland, providing information about the study to different communities, work places, religious groups, education centres and so on, to promote interest and participation from a wider demographic. Also, although the intention was to match patients one-to-one with controls, according to variables such as gender, education and SES, many of the control subjects were female, while the patients were mostly male. Although this was discussed when designing the study, there was also awareness that collecting a control sample within this population could be difficult. Also, in the literature, gender has been a less prominent factor affecting CET performance. It was decided that we would ask the patient with brain injury to nominate who they felt would be happy to complete the test, to increase the chances of finding a control for each patient. It was also noted that more female friends and family visited the unit and also that the female controls were more likely to be willing to participate and complete the questionnaires. In future studies however, it might be helpful to ask people to nominate a male friend or family member initially, and then suggest others if this is not possible. Most patients were, however, matched well according to education and SES, which was important given their strong links to performance on the test in previous studies (O'Carroll et al., 1994, Axelrod & Millis, 1994, MacPherson et al., 2014). Some studies however, have not found associations between education and the CET (e.g. Levinoff et al., 2006). The reason

family and friends of patients were asked initially, was because it was believed more likely that they would better match patients for these important factors. The benefit of this decision in this study was strengthened by the resulting relative homogeneity of the normative sample with regards to education levels and SES. Unfortunately, a fully matched control sample based on friends and family members could not be gathered, and therefore, seven datasets were taken from the normative sample.

Significant differences were not found between the two brain injury groups. This may have been a result of how patients were categorised. It could be possible that the clinicians were not accurate in their decision about individual patients' capacity and the groups were not adequately split. For these complex decisions, there is often a degree of clinical judgement that must be applied. However, as described above, categorisation was consistent and conducted in a systematic, professional manner with the aid of neuropsychological measures, legislation criteria and clinical judgement. A possible explanation of the results is that whilst the processes required for estimation are relevant to making decisions in everyday life, it is possible that a range of other cognitive processes are also involved and if these are impaired they may impact on capacity to make decisions but not on performance on a CET.

10 Sub-test Items

Ten items were identified from two group comparisons that appeared to show sensitivity to the brain injury sample. There was reasonable consistency between the two group comparisons regarding the questions that showed significant differences between the groups, one caveat being that seven participants in the brain injury control group were also in the normative sample. Four of the selected sub-items were related to quantity. It is possible that some of the quantity items may have been perceived as being more challenging, resulting in a stronger feeling of not knowing and therefore more

impulsive estimations. Working memory is important for questions requiring calculations, which may have been useful for some of these questions (e.g. how many times on average do we blink in an hour). Therefore, working memory deficits may also have contributed to larger deviations in these questions. Other than that, it is not clear what might differentiate the questions that appear most sensitive from the others. Most of these questions clearly fit well the key requirements of a cognitive estimation task of being a question to which nobody would be expected to know an exact answer, but a question that, with some problem- solving, a reasonable estimate could be derived. However that might be said of many of the other questions.

Clinical Implications and Recommendations for Future Research

This research has given rise to a new set of items that appear to be sensitive to brain injury. However at this stage the longer question set does not appear reliable or valid enough for clinical use in this population. It may be understandable to conclude on the basis of these results that the CET should be abandoned as an assessment measure. However, there has been a recent interest in the development of new and more clinically rigorous CET questions (e.g. MacPherson et al., 2014) and the present study has provided a clear rationale for the development of a CET, which emphasises fairness across cultures and taps into processes that may show more sensitivity to the deficits in those with brain injury. Future research should attempt to build on these questions to further explore those sensitivity factors, as well as gathering a more heterogeneous normative sample on which to develop a sound scoring method. The items also need to be tested in other clinical samples to further explore whether they might be useful in a clinical context, especially in relation to everyday functioning. Furthermore, as a future research idea, it may be interesting to qualitatively examine how people work through

problems and reach answers that are either within the normal range, or statistically out with the norm. This may provide useful information about the specific barriers to generating reasonable answers, or indeed, the specific difficulties encountered when making a decision about something. CETs could provide a structured, systematic way of determining an individual's needs in rehabilitation settings, and help with personalised goal management training, for example.

Conclusions

Based on the results from this study (and much of the previous literature), cognitive estimation tests do not appear to be reliable or valid enough for use in clinical assessments at the present time. Although this study developed items sensitive to this clinical population, it is important that work continues to explore the use and purpose of this question sub-set in clinical practice, its sensitivity in different neurological populations, and how it relates to individuals' abilities to make decisions in their everyday lives.

References

Axelrod, B.N., & Millis, S.R. (1994). Preliminary standardization of the Cognitive Estimation Test. *Assessment*, **1**(3), 269-274. doi: 10.1177/107319119400100307

Blanchette, I., & Richards, A. (2010). The influence of affect on higher-level cognition: A review of research on interpretation, judgment, decision making and reasoning. *Cognition and Emotion*, **24**(4), 561-595. **doi:**10.1080/02699930903132496

Bland, J. M. & Altman, D. G. (1986). 'Statistical methods for assessing agreement between two methods of clinical measurement', *The Lancet **i***, 307-310.

Brand, M., Kalbe, E., Fujiwara, E., Huber, M., & Markowitsch, H.J. (2003). Cognitive estimation in patients with probable Alzheimer's disease and alcoholic Korsakoff patients. *Neuropsychologica*, **41**(5), 575-584. Retrieved from http://www.sciencedirect.com/science/article/pii/S0028393202001835#

Bullard, S.E., Fein, D, Gleeson, M.K, Tischer, N, Mapou, R.L., & Kaplan, E. (2004) The Biber Cognitive Estimation Test. *Archives of Clinical Neuropsychology*, **19**(6), 835-846. doi:10.1016/j.acn.2003.12.002

Burgess, P.W., Alderman, N., Evans, J., Emslie, H., & Wilson, B.A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society,* **4**(6), 547-558. doi: 10.1017/S1355617798466037

Burgess, P.W., Alderman, N., Evans, J.J., Emslie, H. (1996). Validity of the battery: Relationship between performance on the BADS and ratings of executive problems. In B.A. Wilson (Ed.), *BADS: Behavioural assessment of the Dysexecutive syndrome manual* (pp.18-19). Bury St. Edmunds, UK: Thames Valley Test Company.

Della Sala, S., MacPherson, S.E., Phillips, L.H., Sacco, L., Spinnler, H. (2003). How many camels are there in Italy? Cognitive estimates standardised on the Italian population. *Neurological Science*, **24**(1), 10–15. doi: 10.1007/s100720300015

Department of Health. (2005). *Mental Capacity Act*. London: HMSO.

judgment. 2013. In *Merriam-Webster.com*. Retrieved May, 8 2013, from http://www.merriam-webster.com/dictionary/judgment

Levinoff, E.J., Verret, L., Akerib, V., Phillips, N.A., Babins, L., Kelner, N., & Chertkow, H. (2006). Cognitive estimation impairment in Alzheimer's disease and mild cognitive impairment. *Neuropsychology*, **20**(1), 123-132. doi: 10.1037/0894-4105.20.1.123

MacPherson, S.E., Wagner, G.P., Murphy, P., Bozalli, M., Cipolotti, L., & Shallice, T. (2014). Bringing the Cognitive Estimation Task into the 21st Century: Normative Data on Two New Parallel Forms. *PloS ONE*, **9**(7). doi: 10.1371/journal.pone.0092554

Mariani, C., Sacco, L., Spinnler, H., Venneri, A. (2002). General knowledge of the world: a standardised assessment. *Neurological Sciences*, **23**(4), 161–175. doi: 10.1007/s100720200057

O'Carroll, R., Egan, V., & MacKenzie, D.M. (1994). Assessing cognitive estimation. *British Journal of Clinical Psychology*, **33**(2), 193-197. doi: 10.1111/j.2044-8260.1994.tb01110.x

Office for National Statistics: National Statistics Socio-economic Classification Volume 3 (2010).

Roth, R.M., Pixley, H.S., Kruck, C.L., & Garlinghouse, M.A. (2012). Performance on the cognitive estimation test in schizophrenia. *Applied Neuropsychology: Adult*, **19**(2), 141-146. doi: 10.1080/09084282.2011.595461.

Scottish Government (2000). *Adults with Incapacity (Scotland) Act 2000* (asp 4).

Simblett, S.K., & Bateman, A. (2011), "Dimensions of the Dysexecutive Questionnaire (DEX) examined using Rasch analysis." *Neuropsychological Rehabilitation,* **21**(1), 1-25. doi: 10.1080/09602011.2010.531216

Shallice, T. & Evans, M.E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex*, **14**(2), 294-303. doi:10.1016/S0010-9452(78)80055-0

Silverberg, N.D., Hanks, R.A., & McKay, C. (2007). Cognitive estimation in traumatic brain injury. *Journal of the International Neuropsychological Society*, **13**(5), 898-902. doi: http://dx.doi.org/10.1017/S1355617707071135

Taylor, R., & O'Carroll, R. (1995). Cognitive estimation in neurological disorders. *British Journal of Clinical Psychology*, **34**(2), 223-228. doi: 10.1111/j.2044-8260.1995.tb01456.x

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet*, **2**(7872), 81-4.

Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H.E., & Evans, J.J. (1996). *Behavioural Assessment of the Dyexecutive Syndrome (BADS)*. Bury St Edmunds, UK. Thames Valley Test Company.

Wechsler, D. (2009). *Test of Premorbid Functioning*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008a). *Manual for the Wechsler Assessment of Intelligence Scale-Fourth Ed*. San-Antonio, TX: Pearson.

Weschler, D. (2009). *WMS-IV Technical and Interpretative Manual- Fourth Ed*. San Antonio: Pearson.

Wood, R. LI., (2001). In Wood, R. and McMillan, T. (Eds) *Neurobehavioural Disability and Social Handicap Following Traumatic Brain Injury.* Hove: Psychology Press.

Chapter Three: Advanced Clinical Practice I

Reflective Account

A Reflection on Communication and our Role

Author: Cathy Tran [1]

[1] Mental Health and Wellbeing, University of Glasgow

Correspondence Address:
Mental Health and Wellbeing University of Glasgow
Gartnaval Royal Hospital
1055 Great Western Road Glasgow
G12 0XH
Email: c.tran.1@research.gla.ac.uk

Abstract

My reflective account will focus on my learning experiences throughout training that have been critical to my professional development, with particular reference to communication and our role as Clinical Psychologists. The Integrated Developmental Model (Stoltenberg & Delworth, 1987) is used to structure the main reflection while also drawing upon other reflective models where appropriate. My skills in communication and my confidence in communicating psychological knowledge have grown through a multitude of experiences in training. Certainly, I have faced many challenges, and felt many different emotions at different points. My self-awareness, level of autonomy and motivation has fluctuated, and I have doubted my abilities on many occasions. My main areas of development include effectively communicating with clients, staff and others involved in a client's care, my insight into our role in communication with clients and being aware when I am not communicating in the most helpful way. I hope to show that the challenging times in training are how I have gained confidence and skill but also hope to reflect clearly on my continued learning needs and professional development. As well as communicating on an individual therapeutic level, I feel I have expanded my interest in communicating the need for psychology on a wider level. My understanding of the wider role of a Clinical Psychologist, and my ability, now, to view our profession within a political and public health context, definitely makes me proud to be doing what I am. Personal reflection allows me to continue to improve my skills and gain a wider awareness within our profession. Even my ability to communicate in a reflective manner, and appreciate the importance of reflection has, I hope, developed, and a meta-reflection on this reflective process will conclude the account.

Chapter Four: Advanced Clinical Practice II

Reflective Account

A Trainee's Journey through Research

Author: Cathy Tran [1]

[1] Mental Health and Wellbeing, University of Glasgow

Correspondence Address:
Mental Health and Wellbeing University of Glasgow
Gartnaval Royal Hospital
1055 Great Western Road Glasgow
G12 0XH
Email: c.tran.1@research.gla.ac.uk

Abstract

My reflective account will focus on my research learning experiences throughout training while drawing upon issues in service development and ethical issues in relation to these. The Integrated Developmental Model is used to structure the main body of the reflection and I will also draw upon Gibb's reflective model and the Driscoll reflective model. As I have progressed through training, my awareness of the importance and relevance of research and service development has changed hugely. My main areas of development include a broad range of research skills as well as a curiosity and interest in how research is developed and conducted in practice. My aim is to demonstrate throughout this reflective account how supervision and other experiences have both led me to feel frustrated, but also broadened my interest and increased my enthusiasm. I have seen first hand some of the barriers to conducting research in clinical practice, and have seen the level of detail and precision needed to conduct a project from start to finish. I think that reflecting on this topic would have seemed somewhat pointless to me before I started the course, so the fact I have chosen and been able to communicate my learning on research, service development and ethics for this piece of work, demonstrates the change in my thinking over the past three years. This reflection will be concluded with a meta-reflection on this process and highlight areas of personal professional development.

Appendix 1 Instructions to authors

Journal of the International Neuropsychological Society

Manuscript Length

Research

Regular Research Article: Maximum of 5,000 words (not including abstract, tables, figures, or references) and a 250 word abstract. Regular Research Articles are original, creative, high quality papers covering all areas of neuropsychology; focus may be experimental, applied or clinical.

Systematic Review

Critical Review: Maximum of 7,000 words (not including abstract, tables, figures, or references) and a 250 word abstract. Critical Reviews will be considered on any important topic in neuropsychology. Quantitative meta-analyses are encouraged. Critical Reviews must be preapproved by the Editor-in-Chief. For consideration, please e-mail your abstract to jins@cambridge.org.

General

Manuscript Preparation and Style

The entire manuscript should be typed double-spaced throughout using a word processing program. Unless otherwise specified, the guideline for preparation of manuscripts is the Publication Manual of the American Psychological Association (6th edition) except for references with 3 or more authors (see References section). This manual may be ordered from: APA Order Dept., 750 1st St. NE, Washington, DC 20002-4242, USA.

Pages should be numbered sequentially beginning with the Title Page. The Title Page should contain the full title of the manuscript, the full names and institutional affiliations of all authors; mailing address, telephone and fax numbers, and e-mail address for the corresponding author; and the word count for the abstract and manuscript text (excluding title page, abstract, references, tables, and figures). At the top right provide a short title of up to 45 characters preceded by the lead author's last name. Example: Smith-Memory in Parkinson's Disease. This running head should be repeated at the top right of every following page.

Page 2 should include an Abstract and a list of at least six keywords or mesh terms. Note: structured abstracts must be included with papers submitted after January 1, 2014. A structured abstract must include four header labels: Objective, Method, Results, and Conclusions. A total of six mesh terms (http://www.nlm.nih.gov/mesh/) or keywords should be provided and should not duplicate words in the title.

The full text of the manuscript should begin on page 3. For scientific articles, including Regular Research Articles, Brief Communications, Rapid Communications, and Symposia, the format should include a structured Abstract, Introduction, Method, Results, and Discussion. This should be followed by Acknowledgments, References, Tables, Figure Legends, Figures, and optional Appendices and Supplemental Material.

The use of abbreviations, except those that are widely used, is strongly discouraged. They should be used only if they contribute to better comprehension of the manuscript. Acronyms should be spelled out at first mention. Metric system (SI) units should be used.

Appendices and Supplemental Materials may be submitted. Appendices include material intended for print and should be included with the manuscript file. Supplementary material will appear only online and should be submitted as a separate file.

The Acknowledgements Section should include a disclosure of conflicts of interest (see above) and all sources of financial support for the paper. In documenting financial support, please provide details of the sources of financial support for all authors, including grant numbers. For example, "This work was supported by the National Institutes of Health (grant number XXXXXXX)". Multiple grant numbers should be separated by a comma and space and where research was funded by more than one agency, the different agencies should be separated by a semi- colon with "and" before the final funding agency.

Grants held by different authors should be identified using the authors' initials. For example, "This work was supported by the Wellcome Trust (A.B., grant numbers XXXX, YYYY), (C.D., grant number ZZZZ); the Natural Environment Research Council (E.F., grant number FFFF); and the National Institutes of Health (A.B., grant number GGGG), (E.F., grant number HHHH)."

Tables and Figures

Tables and Figures should be numbered in Arabic numerals. Figures should be numbered consecutively as they appear in the text. Figures should be twice their intended final size and authors should do their best to construct figures with notation and data points of sufficient size to permit legible photo reduction to one column of a two-column format.

Please upload figure(s) in either a .doc or .pdf format. There is no additional cost for publishing color figures. When uploading figures (color or black and white) they need only be a high enough resolution for the reviewers and editors to identify the information you are trying to convey.

The approximate position of each table and figure should be provided in the manuscript: [INSERT TABLE 1 HERE]. Tables and figures should be on separate pages. Tables should have short titles and all figure legends should be on separate pages.

References

References should be consistent with the Publication Manual of the American Psychological Association (6th Edition). In-text references should be cited as follows:

"…Given the critical role of the prefrontal cortex (PFC) in working memory (Cohen et al., 1997; Goldman-Rakic, 1987; Perlstein et al., 2003a, 2003b)…" with multiple references in alphabetical order. Another example: "…Cohen et al. (1994, 1997), Braver et al. (1997), and Jonides and Smith (1997) demonstrated…" References cited in the text with two authors should list both names. References cited in the text with three, four, or five authors, list all authors at first mention; with subsequent citations, include only the first author's last name followed by et al. References cited in the text with six or more authors should list the first author et al. throughout. In the reference section, for works with up to seven authors, list all authors. For eight authors or more, list the first six, then ellipses followed by the last author's name.

Online/Electronic Journal Article with DOI:
Dikmen, S., Machamer, J., Fann, J. & Temkin, N. (2010). Rates of symptom reporting following traumatic brain injury. Journal of the International Neuropsychological Society, 16, 401–411. doi:10.1017/ S1355617710000196

Scientific Article:
Giovannetti, T., Britnell, P., Brennan, l., Siderowf, A., Grossman, M., Libon, D.J., Seidel, G.A. (2012). Everyday action impairment in Parkinson's disease dementia. Journal of the International Neuropsychological Society, 18, 787–798.

Book:
Lezak, M.D., Howieson, D.B., Bigler, E.D., Tranel, D. (2012). Neuropsychological Assessment. New York: Oxford University Press.

Book Chapter:
Mahone, E.M. & Slomine, B.S. (2008). Neurodevelopmental disorders. In J.E. Morgan, & J.H. Ricker (Eds.), Textbook of Clinical Neuropsychology (pp. 105–127). New York: Taylor & Francis.

Report at a Scientific Meeting:
Weintraub, S. (2012, June). Profiles of dementia: Neuropsychological, neuroanatomical and neuropatho- logic phenotypes. International Neuropsychological Society, Oslo, Norway.

Manual, Diagnostic Scheme, etc.:
American Psychiatric Association (1994). Diagnostic and Statistical Manual of Mental Disorders (4th ed.). Washington, DC: American Psychiatric Association Press.

English Editing The Research and Editing Consulting Program (RECP) within the International Neuropsychological Society's International Liaison Committee is designed to provide English language editing and statistical consulting to international colleagues who wish to publish their research in English language journals. For additional information see http://www.the-ins.org/the-research-and-editing-consulting-program.

Proofs The publisher reserves the right to copyedit manuscripts. The corresponding author will receive PDFs for final proofreading. These should be checked and corrections returned within 2 days of receipt. The publisher reserves the right to charge authors for excessive corrections.

Offprints and PDF Files The corresponding author will receive a free pdf. This pdf can also be mounted on the authors' web pages. Offprints must be ordered when page proofs are returned. The offprint order form with the price list will be sent with your PDF.

Open Access Papers In consideration of payment of the Open Access fee specified by Cambridge University Press, the contribution will be published in the Journal of the International Neuropsychological Society within an Open Access environment, freely accessible to those who wish to browse, read, print, save, copy, display or further disseminate the contribution. Please see the Open Access Transfer of Copyright Agreement for the proper procedures at http://journals.cambridge.org/action/displayMoreInfo? Jid=INS&type=tcr. The processes will depend on your source of funding, permissions to use material owned by an outside source, etc.

Systematic Review

Appendix 2.1

Quality Rating Criteria: STROBE-22 Statement for reporting case-controls

Table 1. **The STROBE Statement: a checklist of items that should be addressed in reports of observational studies**

Item	Item number	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any pre-specified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up and data collection
Participants	6	(a) *Cohort study* – Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* – Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* – Give the eligibility criteria, and the sources and methods of selection of participants
		(b) *Cohort study* – For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* – For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders and effect modifiers. Give diagnostic criteria, if applicable
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding
		(b) Describe any methods used to examine subgroups and interactions
		(c) Explain how missing data were addressed
		(d) *Cohort study* – If applicable, explain how loss to follow-up was addressed *Case-control study* – If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* – If applicable, describe analytical methods taking account of sampling strategy
		(e) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report the numbers of individuals at each stage of the study – e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up and analyzed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate the number of participants with missing data for each variable of interest
		(c) *Cohort study* – Summarize follow-up time (e.g. average and total amount)
Outcome data	15*	*Cohort study* – Report numbers of outcome events or summary measures over time *Case-control study* – Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* – Report numbers of outcome events or summary measures

(Table 1, cont.)

Item	Item number	Recommendation
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(b) Report category boundaries when continuous variables were categorized
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done – e.g. analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarize key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalizability	21	Discuss the generalizability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

[a] Give such information separately for cases and controls in case-control studies, and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

An *Explanation and Elaboration* article [18–20] discusses each checklist item, and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the web sites of *PLoS Medicine*, *Annals of Internal Medicine* and *Epidemiology*). Separate versions of the checklist for cohort, case-control and cross-sectional studies are available on the STROBE web site.

Appendix 2.2

Methodological Quality rating

1. **Was there an appropriate and clearly focused question?**

2. **Were patients compared to establish similarities/differences at Baseline and if not, were appropriate measures taken to control for confounding factors?** E.g. SES, level of education/IQ, gender

3. **Was the inclusion/exclusion criteria appropriate?** i.e. did they exclude those with significant confounding factors (e.g. history of psychiatric treatment suggesting disorders with potential neurobehavioural effects, dysphasia, any previous head injury requiring medical treatment, previous episode of unconsciousness, alcoholism, evidence of any neurological disorder (including stroke, seizures, tumours), any chronic medical conditions such as HIV, or first language not English.

4. **Did the study have the power to detect at least a medium effect size (d= 0.5) (51 in each group, if full sample >102) +/- 2.**

Medium effect size d=0.5 (difference of means)- this would be a clinically useful cut-off. E.g. if effect size found is very small, it is unlikely the power is sufficient.

For this quality question:
effect size = 0.5
Power = 0.8
Alpha = 0.5
Therefore looking for **51 per group** for it to be sufficient.

5. **If any other tests were used were they valid and reliable? Was the rationale for other tests used clear?**

6. **Was a statistically valid approach used to define 'bizarre answers' to questions?** *(An approach that is statistically valid would need to have answers that are compared to a sample of healthy individuals/normed on healthy group, and a cut-off for impaired performance suggested based on variability of answers).*

7. **Cases are clearly defined and differentiated from controls. (from SIGN case-control)**

8. **It is clearly established that controls are non-cases** (e.g. BI out with frontal area?) **(from SIGN case-control)**

Appendix 2.3

Rater Scorings

Original Score disparities and final scores in format: e.g. 10 *Reviewer 1* (11 *Reviewer 2*) = 11 *final score*

Appolonio et al. (2003)	***5 R1 (10 R2) = 4 final***
Barrera et al. (2005)	***13 (9) = 11***
Brand et al. (2003a)	***13 (11) = 12***
Brand et al. (2003b)	***11(13) = 10***
Bullard et al. (2004)	***11 (12) = 9***
Burgess et al. (1998)	***14 (13) = 12***
Dixon et al. (2004)	***15 (15) = 14***
Kopelman (1991)	***13 (11)= 13***
Leng&Parkin (1988)	***5 (4) = 5***
Levinoff et al. (2006)	***13 (17) = 13***
Manning et al. (2005).	***15 (9) = 11***
Mendez et al. (1998)	***11 (16) = 10***
Nedjam et al. (2004)	***9 (9) = 10***
Parente et al. (2013)	***10(9) = 11***
Roth et al. (2012)	***13 (12) = 14***
Silverberg et al. (2007)	***14 (15) = 14***
Shallice & Evans (1978)	***8 (13) = 9***
Shoqueirat et al. (1990)	***9 (8) = 10***
Spencer & Johnson-Green (2009)	***14 (14) = 13***
Taylor & O'Carroll.	***11 (10) = 9***
Treitz et al. (2009)	***12 (13) = 12***

Major Research Project

Appendix 3.1

Letters of Ethical Approval and Insurance

University of Glasgow | College of Medical, Veterinary & Life Sciences

14th November 2014

Dear Cathy Tran, Dr Brian O'Neill, Professor Jonathan Evans«Principal_Investigator»

MVLS College Ethics Committee

Project Title: Developing a Culture Fair Cognitive Estimates Test
Project No: 200140002

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. They are happy therefore to approve the project, subject to the following conditions

- Project end date: November 2015
- The research should be carried out only on the sites, and/or with the groups defined in the application.
- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.
- You should submit a short end of study report to the Ethics Committee within 3 months of completion.

Yours sincerely

[redacted]

Prof. Andrew C. Rankin
Deputy Chair, College Ethics Committee

Andrew C. Rankin
Professor of Medical Cardiology
BHF Glasgow Cardiovascular Research Centre
College of Medical, Veterinary & Life Sciences
University of Glasgow, G12 8TA
Tel: 0141 211 4833
Email: andrew.rankin@glasgow.ac.uk

Heritage Insurance
Solutions Limited
1 Cornhill
London
EC3V 3ND

Tel: +44 (0) 207 220 9020
Fax: +44 (0) 203 014 7695

info@heritage.co.gg

www.heritage.co.gg

1st July 2015

To Whom it May Concern

Dear Sirs

Confirmation of Insurance Cover

NAME OF ORGANISATION	The Disabilities Trust Limited known as The Disabilities Trust
SERVICE DESCRIPTION	The Insurances in place cover all the activities of the DT Group.
EMPLOYERS LIABILITY:	
(a) Insurers (name & address)	Barbican Holdings (UK) Limited (Syndicate 1955), Lloyds UMR: 86053BCI2003 Registered office – 33 Creechurch Lane, London, EC3A 5EB
(b) Policy Number	BCI08001/689
(c) Renewal Due Date	Next Renewal date – 01/07/2015
(d) Does the policy include	
(1) Liability assumed under contract	Yes
(2) Principal Indemnity Clause	Yes
(e) Limit of Indemnity	£10,000,000
(f) Details of any restrictive endorsements or Warranties	Nothing specific to the client other than one would expect as standard on an EL policy.
PUBLIC LIABILITY	
(a) Insurers (name & address)	Barbican Holdings (UK) Limited (Syndicate 1955), Lloyds UMR: 86053BCI2003 Registered office – 33 Creechurch Lane, London, EC3A 5EB
(b) Policy Number	BCI08001/689
(c) Renewal Due Date	Next Renewal date – 01/07/2015
(d) Does the policy include	
(1) Liability assumed under contract?	Yes
(2) Principal Indemnity Clause? (3) Indemnity in respect of administration of prescribed medicines?	Yes Yes
(e) Limit of Liability	
(1) What is the limit on any one accident?	£10,000,000 - £5,000,000 provided by Barbican and £5 mil over £5 mil provided by ACE European Group Limited @ ACE Building, 100 Leadenhall Street, London, EC3A 3BP.

(2) Is there a lower limit for certain risk?	Treatment Extension limited to £2,000,000; however Full Medical Malpractice Insurance in place up to £5,000,000 with Medical Professional Liability Company Limited @ 107 Fenchurch Street, London, EC3M 5JF
(3) Is there an aggregate limit?	Limit is aggregate on the Med Mal. Barbican for Public Liability – No aggregate
(f) Details of Excess	£500 third party property damage
(g) Details of any restrictive endorsements or warranties	Other than standard endorsements / warranties the only specific one relates to off premises fund raising events where the more hazardous events need referral.
PROFESSIONAL INDEMNITY	
(a) Insurers (name & address)	Markel (London) Limited @ The Markel Building, 49 Leadenhall Street, London, EC3A 2EA
(b) Policy Number	ZC1908T120GR/126
(c) Renewal Due Date	Next Renewal date – 01/07/2015
(d) Limit of Liability	
(1) What is the limit on any one claim, act or occurrence?	£2,000,000
(e) Details of Excess	£2,500
(f) Details of any restrictive endorsements or warranties	Nothing other than standard that you would expect to see on a Combined Directors & Officers Liability and Professional Indemnity policy.
MEDICAL MALPRACTICE INSURANCE	
(a) Insurers (name & address)	The Medical Professional Liability Company Limited, 107 Fenchurch Street, London, EC3M 5JF
(b) Policy Number	013/0005893/00
(c) Renewal Due Date	Next Renewal date – 01/07/2015
(d) Limit of Liability	
(1) What is the limit on any one claim, act or occurrence?	£5,000,000
(e) Details of Excess	£2,500
(f) Details of any restrictive endorsements or warranties	Nothing other than standard that you would expect to see on a Medical Malpractice policy.

We confirm that the insurance arrangements of the organisation named above are as set out above and the premiums have been paid.

K D Bradley

Karl Bradley, Managing Director on behalf of Heritage Insurance Solutions Limited

Appendix 3.2

Health and Safety Forms

WEST OF SCOTLAND/ UNIVERSITY OF GLASGOW
DOCTORATE IN CLINICAL PSYCHOLOGY

HEALTH AND SAFETY FOR RESEARCHERS

1. Title of Project	Developing a Culture Fair Cognitive Estimates Test
2. Trainee	Cathy Tran
3. University Supervisor	Professor Jon Evans
4. Other Supervisor(s)	Dr. Brian O'Neill
5. Local Lead Clinician	Dr. Brian O'Neill
6. Participants: (age, group or sub-group, pre- or post-treatment, etc)	Normative group: volunteers from general population, age range 18+ Severe-extremely severe brain injured patients, age range 18+
7. Procedures to be applied (eg, questionnaire, interview, etc)	A normative data sample will be collected by sending questionnaires via Qualtrics (Online database), sending hard copies with free post, and testing participants face to face. Will administer the newly developed cognitive estimates test, a demographic questionnaire and the Dysexecutive Questionnaire (DEX) to patients with brain-injuries.
8. Setting (where will procedures be carried out?) i) General	Brain Injury Rehabilitation Trust, Graham Anderson House, Glasgow. Possible extension to other BIRT units. This is a specialist neurobehavioural assessment and post-acute rehabilitation hospital for people with a non-progressive acquired brain injury.

	The service has its own Health and Safety Policies and Procedures and the following will be ensured: 1. The researcher will have a tour of the building by a member of maintenance staff to demonstrate location of emergency exits. 2. The researcher will receive training on fire procedures. 3. The researcher will demonstrate evidence of Disclosure Scotland clearance to work with vulnerable groups. 4. The researcher will have the use of a personal alarm when working in clinical areas. 5. The researcher will be appraised of risk assessments pertaining to participants. 6. The researcher will be able to contact site supervisor or service manager for discussion of any concerns.
ii) Are home visits involved	N

WEST OF SCOTLAND/ UNIVERSITY OF GLASGOW
DOCTORATE IN CLINICAL PSYCHOLOGY

HEALTH AND SAFETY FOR RESEARCHERS

9. Potential Risk Factors Identified (see chart)	For participants: For those unfamiliar with a cognitive estimates test, the fact that the questions have no exact answer, and may appear impossible, may cause some confusion, a lack of motivation or beliefs of being unable to do the test. This patient group may be particularly vulnerable to experiencing these feelings. For researchers: Researcher will be working with people who have severe brain injuries. Due to the nature of these injuries, unpredictable and aggressive behaviours are possible.
. 10. Actions to minimise risk (refer to 9)	Aim to follow local procedures to ensure staff and patient safety. BIRT has procedures in place to minimise risk to staff and these are thought to be adequate in the context of the proposed study. **Local BIRT procedures to ensure staff safety will be adhered to.** **Researcher Safety:** Clinicians who are familiar with the client will be asked only to refer patients who do not represent a significant risk. The researcher will wear an alarm during visits. **Participant Safety:** The procedures used in this study are similar to those administered by clinicians with these participants and are not usually associated with the onset of significant distress. To avoid distress regarding the questions as much as possible, it will be made clear that the nature of the questions and answers require a reasoned guess and that there is no exact answer.

Trainee signature: ..... Date: .....

University supervisor signature: ..... Date:.....

WEST OF SCOTLAND/ UNIVERSITY OF GLASGOW
DOCTORATE IN CLINICAL PSYCHOLOGY

HEALTH AND SAFETY FOR RESEARCHERS

Points to consider when assessing risk. If any answer is "no" then make a case for the design being safe or reconsider the design of the study.

Participants	
Yes	**No**
This participant sample is not normally associated with dangerous or unpredictable behaviour	This participant sample is associated with impulsive, irrational or unpredictable behaviour, and/or has poor emotional control

Procedures	
Yes	**No**
The procedures in the study are same/similar to those used by clinical psychologists with these participants and are not normally associated with production of significant distress.	These are novel procedures, are not used with this group and by their nature might produce anger, irritability or distress.

Settings	
Yes	**No**
These are clinical or University research settings, or other institutional settings, that participants routinely attend (eg, a school). They have procedures in place to minimise risk to staff and these are thought to be adequate in the context of the proposed study.	A private or other setting where there are not health and safety procedures that are relevant to research or clinical work proceeding without risk

Version 3/10/06

Appendix 3.3

Culture Fair Cognitive Estimation Test

Thank you for agreeing to participate in this study. Please answer the questions below. Please remember that the questions are ones for which there is no absolutely correct answer and it is not expected that anyone would be likely to know an exact answer. This is the case even for the questions where you are asked to estimate a number – we are just interested in your estimates.

1. How many steps does it take for an adult to walk 100m (110 yards/ 330 feet)?	
2. How many TV programs are there, on average, on any one channel between 6 and 11pm?	
3. How long is an adult elephant's trunk?	
4. How many mouthfuls are there in an average bowl or plate of food?	
5. How much does a hen's egg weigh?	
6. How long is an average train carriage?	
7. How many times on average do we blink in an hour?	
8. How much do 12 average sized bananas weigh?	
9. How long does it take an adult to put on clothes in the morning?	
10. How many eyelashes do we have on our lower eyelid?	
11. How fast do lions run?	
12. What is the distance an adult can walk in an afternoon (6hours)?	

13. How many spokes are there on a bicycle wheel?	
14. What quantity of water would fill a football?	
15. How tall is the average woman?	
16. How much does a pair of men's shoes weigh?	
17. How many grains of rice are there in a single handful?	
18. How long would it take to peel a large orange?	
19. How much liquid would you be able to squeeze out of 6 lemons?	
20. What is the length of an average man's spine?	
21. How heavy is a man's cotton shirt?	
22. How long does it take for a young man to walk 8km (5 miles)?	
23. How far could an ox pull a farm cart in one hour?	
24. How many brushings can someone get from an average tube of toothpaste?	
25. How much does a wooden dining chair weigh?	
26. How many feathers are there on an adult hen?	
27. How fast does a swallow fly?	
28. How long would it take an adult to hand write a one-page letter?	
29. How heavy is a full grown elephant?	
30. How many times can you wash your hands from a bar of soap?	

Thank you.

- Questions taken from the BCET questionnaire (Bullard et al. 2004): 12, 16, 24, 25
- Questions taken from the original Shallice and Evans (1978) CET: 2, 15, 20
- Questions taken from Della Sala et al. (2003): 10, 21, 27
- Questions taken from Axelrod & Millis (1994): 29

Appendix 3.4

Demographic Data Form

Version 2 2.12.14

Developing a Culture Fair Cognitive Estimates Test

Demographic Data Form

Age:	
Gender:	
Age Left Education/Still studying:	
Level of Education:	
Occupation	
Right or left handed	
Postcode:	

Appendix 3.5

Participant Consent Forms, Initial Letters and Information Sheets

Patient/family Consent Form

Version 2 29.7.14

Patient/Family Consent Form

Developing a Culture Fair Cognitive Estimates Test

Contact details:
Cathy Tran
2019731t@student.gla.ac.uk

1. I confirm that I have read and understand the participant information leaflet for the above study. ☐
2. I confirm that the researcher has answered any queries to my satisfaction. ☐
3. I understand that my participation is voluntary and that I am free to withdraw from the project at any time, without my medical care or legal rights being affected, and that information I have provided up to that point may be included anonymously in the results of the study. ☐
4. I understand that any information collected about me in the study will remain confidential, and that no information that identifies me will be made publicly available. ☐
5. I understand that my data (including personal information) may be accessed by authorised representatives of the Brain Injury Rehabilitation Trust (the Sponsor) for the purposes of audit only. ☐
6. I agree that my clinical team may be informed of my participation in this study.** ☐
7. I agree that this interview may be audio recorded and may be used in a further study to assess how people make decisions. ☐
8. I consent to being a participant in this study. ☐
9. Please tick this box if you would be willing to be contacted again to complete this questionnaire on a second occasion. ☐

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

Participant Consent Form

Version 2 29.7.14

Participant Consent Form

Developing a Culture Fair Cognitive Estimates Test

Contact details:
Cathy Tran
2019731t@student.gla.ac.uk

1. I confirm that I have read and understand the participant information leaflet for the above study. ☐
2. I confirm that the researcher has answered any queries to my satisfaction. ☐
3. I understand that my participation is voluntary and that I am free to withdraw from the project at any time, and that information I have provided up to that point may be included anonymously in the results of the study. ☐
4. I understand that any information collected about me in the study will remain confidential, and that no information that identifies me will be made publicly available. ☐
5. I consent to being a participant in this study. ☐
6. Please tick this box if you would be willing to be contacted again to complete this questionnaire on a second occasion. ☐

----------	----------	----------
Name of Participant	Date	Signature
----------	----------	----------
Name of Researcher	Date	Signature

Initial letter to Participants

Version 1 09.12.14

Developing a Culture Fair Cognitive Estimates Test

Hello,

We would like to invite you to take part in a study exploring how people make judgments. We are developing a new culture fair test based on questions where few people should know the correct answer but most may be able to estimate.

In this phase of the study, we are inviting healthy people (without history of significant neurological disease, psychiatric disorder or acquired brain injury) to complete the measure as best they can. This will help us to gather some normative data for the measure. In turn, this will help clinicians working with people who have suffered a brain injury to identify people who have difficulty making estimations and judgements. We hope that this may add to the ways of identifying support needs among people with brain impairments. More detailed information is available on the attached information sheet.

We would be very grateful for your participation.

Thank you for your time.

Cathy Tran, Trainee Clinical Psychologist ,
c.tran.1@research.gla.ac.uk

Jonathan Evans, Professor of Applied Neuropsychology,
jonathan.evans@glasgow.ac.uk

Brian O'Neill, Consultant in Neuropsychology and Rehabilitation,
brian.oneill@thedtgroup.org

Initial letter to Patients

Version 2 26.9.14

Developing a Culture Fair Cognitive Estimates Test

Contact details:
Cathy Tran
2019731t@student.gla.ac.uk

Hello,

We would like to invite you to take part in our study. We are exploring how people make judgments and we are developing a new test that we hope will be useful whatever country or culture a person comes from. More details of the study are provided in the Participant Information Sheet.

If you would like to help with our study, you will meet the researcher who will ask you to answer some questions (this should no more than 20-30 minutes of your time). If you would like to know more, please inform a member of the clinical team, who will give you some further information on the study.

Thank you for your time.

Ms Cathy Tran

Professor Jon Evans

Dr Brian O'Neill

Initial letter to friends/family members

Version 2 29.7.14

Developing a Culture Fair Cognitive Estimates Test

Contact details:
Cathy Tran
2019731t@student.gla.ac.uk

Hello,

We would like to invite you to take part in our study. We are exploring how people form judgments and make decisions. Brain injury can affect the ability to form judgements and make decisions, and this can have important consequences. We are developing a new culture fair clinical test that we hope will be able to be used in different countries and cultures to help clinicians assess estimation skills. These skills are important in the process of forming judgements and making decisions. Whilst we are interested in how brain injury affects these abilities, we need to see how people who have not had a brain injury do the task in order to compare with people who have had a brain injury. Further details are provided in the attached Information sheet.

If you are willing to discuss taking part please let a member of the team at Graham Anderson House know or contact Brian O'Neill or Cathy Tran (contact details below). You will then meet the researcher (Cathy Tran) who can answer any questions you have about the project. If you are willing to participate, Cathy will ask you to sign a consent form and then answer some questions (this should take around 20 minutes of your time in total).

Thank you for your time.

Dr. Brian O'Neill,
Consultant in Neuropsychology
and Rehabilitation
Graham Anderson House
1161 Springburn Road
Glasgow
G21 1UU

Phone: 0141 4046060

Email: brian.oneill@thedtgroup.org

Cathy Tran
Trainee Clinical Psychologist
Mental Health and Wellbeing
The Academic Centre
1055 Great Western Road
Glasgow
G12 0XH

Email : c.tran.1@research.gla.ac.uk

Participant Information Sheet

Version 3 26.9.14

Participant Information Sheet

We would like to invite you to take part in our research study. Before you decide whether you would like to take part, we would like you to understand why the research is being done and what it would involve for you.

Please ask if there is anything that is not clear.

Who is conducting the research?

This study is being carried out by researchers from the Brain Injury Rehabilitation Trust and the University of Glasgow. The research team includes Dr Brian O'Neill, Consultant Clinical Psychologist, Professor Jonathan Evans, Professor of Applied Neuropsychology at the University of Glasgow, and Cathy Tran, who is a trainee Clinical Psychologist at the University of Glasgow.

Purpose of the study

We all make judgements in our daily lives. When people suffer an injury to the brain their judgment and decision-making may be impaired. There have been attempts to develop measures of judgement, which have been used to assess these abilities following brain injury. The most common of these are Cognitive Estimate Tests (CETs), where people are asked questions where the answer is difficult to know exactly, and estimation is required. Difficulties in this ability can be caused by many forms of injury to the brain.

A number of estimation tasks have been developed and used in clinical populations. Previous tests include questions that would not be fair for people in all cultures, as they include questions that would be very familiar in one country but not another. We have developed a new cognitive estimation test that we hope will be able to be used in different countries and cultures.

The aim of this study is to see the differences between how those with acquired brain injuries answer these questions compared to those who do not have a brain injury. We then also want to see whether this can be a clinically useful test, by comparing answers between two groups of people with brain injuries- those deemed able to have capacity to make decisions about their welfare, versus those who do not have this capacity.

Why have I been asked to take part?

To determine what answers are typical, we must gather a range of answers from the general population for each estimate question. Only by doing this, will we be able to make any meaningful comparisons between groups. Therefore, we would be grateful if you would take the time to complete the questions, which we estimate will take around 20 minutes of your time.

Do I have to take part?

It is up to you to decide to join the study. We will give you information on the study and if you agree to take part, we will then ask you to consent to taking part. You are free to withdraw at any time, without giving a reason.

What will I be asked to do?

After agreeing to take part, you will be asked a few questions about yourself, for example your age, and how many years of education you have had.

Then you will be asked to answer a set of 30 questions asking you to estimate answers to questions. This should take no more than about 20 minutes. There is no time limit on the questions, and you can take your time on these if you wish. The questions are ones for which there is no absolutely correct answer or it is not expected that anyone would be likely to know an exact answer. This is the case even for the questions where you are asked to estimate a number – we are just interested in your estimates.

If you are being tested face to face, if you agree, then we may audio record this interview. This recorded information will not be used in the current study, however may be used in a future research study exploring and analysing the reasoning and decision making process. This recording would be anonymous.

We may ask if you would be willing to complete the questionnaire on a second occasion, 2-3 weeks after the first time.

What are the possible advantages of taking part?

There will not be major advantages to taking part other than helping with our research. We hope that our research will help develop effective ways to assess whether people with brain injuries may have difficulties making important decisions about their lives.

What are the possible disadvantages of taking part?

We do not foresee any negative effects of taking part. Taking part will probably involve about 20 minutes or less of your time.

What happens to the information collected?

Your identity and personal information will be kept completely confidential and known only to the research team. All information obtained during the study will remain confidential and stored within a locked filing cabinet at Graham Anderson House or on a secure University computer. Information collected will be made anonymous by the use of a unique identity code. Information is held in accordance with the Data Protection Act, which means that we keep it safely and cannot reveal it to other people without your permission. Your information will not be personally identifiable in any reports about the results of the study.

Only three members of the research team will have access to these details.

All research in BIRT is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Glasgow University Research Ethics Committee.

What if I have any further questions?
We will give you a copy of the information sheet and signed consent form to keep. If you would like more information about the study and wish to speak to someone not closely linked to the study, please contact Dr Sue Copstick, Clinical Director for the Brain Injury Rehabilitation Trust (Sue.Copstick@thedtgroup.org).

Thank you for reading this information and for considering taking part in our study.

Research Team Contacts for Further Information

For further information from the research team please contact:

Dr. Brian O'Neill,
Consultant in Neuropsychology
and Rehabilitation
Graham Anderson House
1161 Springburn Road
Glasgow
G21 1UU

Phone: 0141 4046060

Email: brian.oneill@thedtgroup.org

Cathy Tran
Trainee Clinical Psychologist
Mental Health and Wellbeing
The Academic Centre
1055 Great Western Road
Glasgow
G12 0XH

Email : c.tran.1@research.gla.ac.uk

Patient Information Sheet

Version 3 30.9.14

Patient Information Sheet

We would like to invite you to take part in our research study. Before you decide whether you would like to take part, we would like you to understand why the research is being done and what it would involve for you. One of our team can answer any questions you have. Participating in the study should only take about 20 – 30 minutes.

Purpose of the study

We all make judgements in our daily lives. Some judgements are simple with very little chance of a bad outcome (e.g. whether to have tea or coffee at the cafe), but some are very important and could lead to negative consequences if a poor decision is made (e.g. where to live, who to have a relationship with). After brain injury some people find it more difficult to make decisions. This study is looking into judgment and decision-making and how we might be able to assess this ability. One way of assessing this is by asking questions where the answer is difficult to know exactly, and estimates are required.

We have developed a new assessment tool that we would like you to try. It involves estimating the answers to a set of questions. We have also asked many others to complete the questions and we are interested in the differences in how people answer them.

Why have I been asked to take part?

You are being asked to take part because you have had a brain injury. We are asking a group of people who have had a brain injury and a group of people who have not had a brain injury.

We will also be asking others who are staying at Graham Anderson House to complete these answers, as well as asking a member of your family, if this is possible, and if they agree to take part.

Do I have to take part?

It is up to you to decide to take part. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent

form. If you wish, you can take a week to think about it. You are free to withdraw at any time, without giving a reason. This would not affect the standard of care you receive.

What will I be asked to do?

If you agree to take part, Cathy Tran, the researcher, will contact you. After signing a consent form, you will be asked a few questions about yourself, for example your age, and how many years of education you have had. You will then be asked to complete a short questionnaire which asks for your views on any difficulties you feel you have with things like solving problems, planning for the future and making decisions.

Then you will be asked to answer a set of 25 questions involving estimating the answers. There are no exact answers to any of these questions. Even the questions where you are asked to estimate a number do not have an absolutely correct answer- we are interested in your estimates. This should take no more than about 10-15 minutes. There is no time limit on the questions, and you can take your time on these if you wish. The researcher will be with you, so can help if there is anything you are unsure about.

If you agree, then we may audio record this interview. This recorded information will not be used in the current study, however may be used in a future research study exploring and analysing the reasoning and decision making process. This recording would be anonymous.

We may ask if you would be willing to complete the questionnaire on a second occasion, 2-3 weeks after the first time.

What are the possible advantages of taking part?

There will not be major advantages to taking part other than helping with our research. We hope that our research will help develop effective ways to assess whether people with brain injuries may have difficulties making important decisions about their lives.

What are the possible disadvantages of taking part?

We do not foresee any negative effects of taking part. Taking part will probably involve about 20 - 30 minutes or less of your time.

What happens to the information collected?

Your identity and personal information will be kept completely confidential and known only to the research team. All information obtained during the study will remain confidential and stored within a locked filing cabinet at Graham Anderson House or on a secure University computer. Information collected will be made anonymous by the use of a unique identity code. Information is held in accordance with the Data Protection Act, which means that we keep it safely and cannot reveal

it to other people without your permission. Your information will not be personally identifiable in any reports about the results of the study.
Only three members of the research team will have access to these details.

All research in BIRT is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Glasgow University Research Ethics Committee.

What if I have any further questions?
We will give you a copy of the information sheet and signed consent form to keep. If you would like more information about the study and wish to speak to someone not closely linked to the study, please contact Dr Sue Copstick, Clinical Director for the Brain Injury Rehabilitation Trust (Sue.Copstick@thedtgroup.org).

Thank you for reading this information and for considering taking part in our study.

Research Team Contacts for Further Information

For further information from the research team please contact:

Dr. Brian O'Neill,
Consultant in Neuropsychology
and Rehabilitation
Graham Anderson House
1161 Springburn Road
Glasgow
G21 1UU

Phone: 0141 4046060

Email: brian.oneill@thedtgroup.org

Cathy Tran
Trainee Clinical Psychologist
Mental Health and Wellbeing
The Academic Centre
1055 Great Western Road
Glasgow
G12 0XH

Email : c.tran.1@research.gla.ac.uk

Family/friend Information Sheet

Version 3 30.9.14

Family/friend Information Sheet

We would like to invite you to take part in our research study. Before you decide whether you would like to take part, we would like you to understand why the research is being done and what it would involve for you. One of our team can answer any questions you have. Participating in the study should only take about 20 minutes.

Please ask if there is anything that is not clear.

Who is conducting the research?

This study is being carried out by researchers from the Brain Injury Rehabilitation Trust and the University of Glasgow. The research team includes Dr Brian O'Neill, Consultant Clinical Psychologist, Professor Jonathan Evans, Professor of Applied Neuropsychology at the University of Glasgow, and Cathy Tran, who is a trainee Clinical Psychologist at the University of Glasgow.

Purpose of the study

We all make judgements in our daily lives. This study is looking into judgment and decision-making and how we might be able to assess this ability. One way of assessing this is by asking questions where the answer is difficult to know, and estimates are required.

We have developed a new set of questions that we would like you to try and answer. We have also asked many others to complete the questions and we are interested in the differences in how people answer them. We want to see whether there are differences in the way these questions are answered between those who have suffered a brain injury versus those who have not. In addition to this, we would like to compare two groups of people who have acquired brain injuries- those who are considered to have the capacity to make decisions about their welfare versus those who do not have this ability as a result of their brain injury. This will help us to see whether our questionnaire could help clinicians who have to make judgements regarding whether people with brain injury have the capacity to make important decisions.

Why have I been asked to take part?

Whilst we are asking those who have acquired a brain injury to participate, we are also asking people who have not acquired a brain injury. We intend to look at the differences in answers between those who have a brain injury and those who do not. To ensure our comparisons are meaningful, it can be helpful to have a close relative or friend to take part in the study, as it is likely that you will have a number of things in common.

Do I have to take part?

It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason.

We will also be asking other friends and family members of those staying at Graham Anderson House to complete these answers.

What will I be asked to do?

If you agree to take part, Cathy Tran, the researcher, will contact you. Firstly, you will be asked to answer a few questions about yourself, for example your age, and how many years of education you have had. You will then be asked to complete a questionnaire about how _____ makes decisions in his/her daily life.

Thirdly, you will be asked to answer a set of 25 questions asking you to estimate answers to questions. This should take around 20 minutes in total. There is no time limit on the questions, and you can take your time on these if you wish. The questions are ones for which there is no absolutely correct answer. This is the case even for the questions where you are asked to estimate a number – we are just interested in your estimates.

The researcher will be with you, so they will be there if there is anything you are unsure about.

If you agree, then we may audio record this interview. This recorded information will not be used in the current study, however may be used in a future research study exploring and analysing the reasoning and decision making process. This recording would be anonymous.

We may ask if you would be willing to complete the questionnaire on a second occasion, 2-3 weeks after the first time.

What are the possible advantages of taking part?

There will not be major advantages to taking part other than helping with our research. We hope that our research will help develop effective ways to assess whether people with brain injuries may have difficulties making important decisions about their lives.

What are the possible disadvantages of taking part?

We do not foresee any negative effects of taking part. Taking part will probably involve about 20 minutes or less of your time.

What happens to the information collected?

Your identity and personal information will be kept completely confidential and known only to the research team. All information obtained during the study will remain confidential and stored within a locked filing cabinet at Graham Anderson House or on a secure University computer. Information collected will be made anonymous by the use of a unique identity code. Information is held in accordance with the Data Protection Act, which means that we keep it safely and cannot reveal it to other people without your permission. Your information will not be personally identifiable in any reports about the results of the study.

Only three members of the research team will have access to these details.

All research in BIRT is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Glasgow University Research Ethics Committee.

What if I have any further questions?

We will give you a copy of the information sheet and signed consent form to keep. If you would like more information about the study and wish to speak to someone not closely linked to the study, please contact Dr Sue Copstick, Clinical Director for the Brain Injury Rehabilitation Trust (Sue.Copstick@thedtgroup.org).

Thank you for reading this information and for considering taking part in our study.

Research Team Contacts for Further Information

For further information from the research team please contact:

Dr. Brian O'Neill,
Consultant in Neuropsychology
and Rehabilitation
Graham Anderson House
1161 Springburn Road
Glasgow
G21 1UU

Phone: 0141 4046060

Email: brian.oneill@thedtgroup.org

Cathy Tran
Trainee Clinical Psychologist
Mental Health and Wellbeing
The Academic Centre
1055 Great Western Road
Glasgow
G12 0XH

Email : c.tran.1@research.gla.ac.uk

Appendix 3.6

CFACE Error Scoring Method

CFACE error scoring method

Q1	Percentile	Answers	Error Score given
	26-75%	107-190	0
	21-25% 76-80%	101-106 191-200	1
	16-20% 81-85%	99-100 201-225	2
	11-15% 86-90%	69-98 226-259	3
	91-95% 5-10%	260-314 45-68	4
	>95% <5%	315 and above 44 and under	5

Q2	Percentile	Answers	Error Score given
	26-75%	5.6-7.7	0
	76-80% 21-25%	7.8-7.9 5.1-5.5	1
	81-85% 16-20%	8-8.5 4.9-5.0	2
	86-90% 11-15%	8.6-9.0 4.7-4.8	3
	91-95% 5-10%	9.1-9.5 4-4.6	4
	>95% <5%	9.6 and above <4	5

Q3	Percentile	Answers	Error Score given
	26-75%	143-214	0
	76-80% 21-25%	215-245 123-142	1
	81-85% 16-20%	246-280 121-122	2
	86-90% 11-15%	281-300 100-120	3
	91-95% 5-10%	301-400 86-99	4
	>95% <5%	401 and above 85 and below	5

Q4	Percentile	Answers	Error Score given
	26-75%	16-28	0
	76-80% 21-25%	29 15	1
	81-85% 16-20%	30-32 14.0	2
	86-90% 11-15%	33-38 13.0-13.9	3
	91-95% 5-10%	39-48 10-12	4
	>95% <5%	49 and above 9 or below	5

Q5	Percentile	Answers	Error Score given
	26-75%	28.4-95	0
	76-80% 21-25%	96-113.0 22-28.3	1
	81-85% 16-20%	113.1-150 17-21	2
	86-90% 11-15%	151- 227 7-16	3
	91-95% 5-10%	228-460 6-11	4
	>95% <5%	461 and above 5 and below	5

Q6	Percentile	Answers	Error Score given
	26-75%	11.1-27	0
	76-80% 21-25%	27.1- 29 9.7-11	1
	81-85% 16-20%	29.1-45 9.1- 9.6	2
	86-90% 11-15%	45.1-81 6.2-9	3
	91-95% 5-10%	81.1-149 5-6.1	4
	>95% <5%	150 and above 4.9 and below	5

Q7	Percentile	Answers	Error Score given
	26-75%	191-1150	0
	76-80% 21-25%	1151-1200 151-190	1
	81-85% 16-20%	1201-1600 100-150	2
	86-90% 11-15%	1601-2399 60-99	3
	91-95% 5-10%	2400-2999 20-59	4
	>95% <5%	3000 and above 19 and below	5

Q8	Percentile	Answers	Error Score given
	26-75%	907-1850	0
	76-80% 21-25%	1851-1999 901-906	1
	81-85% 16-20%	2000-2268 500-900	2
	86-90% 11-15%	2269-2799 411-499	3
	91-95% 5-10%	2800-3600 123-410	4
	>95% <5%	3601 and above 122 and below	5

Q9	Percentile	Answers	Error Score given
	26-75%	120-255	0
	76-80% 21-25%	256-270 110-119	1
	81-85% 16-20%	271-360 100-109	2
	86-90% 11-15%	361-465 90-99	3
	91-95% 5-10%	466-599 56-89	4
	>95% <5%	600 and above 55 and below	5

Q10	Percentile	Answers	Error Score given
	26-75%	25-59	0
	76-80% 21-25%	60-68 20-24	1
	81-85% 16-20%	69-98 19.6-19.9	2
	86-90% 11-15%	99-100 19-19.5	3
	91-95% 5-10%	101- 180 13-18	4
	>95% <5%	181 and above 12 and below	5

Q11	Percentile	Answers	Error Score given
	26-75%	25-43.4	0
	76-80% 21-25%	43.5-49.7 24.9	1
	81-85% 16-20%	49.8 19.1-24.8	2
	86-90% 11-15%	49.9 18.7-19.0	3
	91-95% 5-10%	50-59 14-18.6	4
	>95% <5%	60 and above 13 and below	5

Q12	Percentile	Answers	Error Score given
	26-75%	20.9-37	0
	76-80% 21-25%	37.1-38 19.1-20.8	1
	81-85% 16-20%	38.1-40.2 16.2-19	2
	86-90% 11-15%	40.3-47 14-16.1	3
	91-95% 5-10%	47.1-48 8-13	4
	>95% <5%	48.1 and above 7 and below	5

Q13	Percentile	Answers	Error Score given
	26-75%	20.2-45	0
	76-80% 21-25%	45.1-49 19.1-20.1	1
	81-85% 16-20%	49.1-49.9 18.1-19	2
	86-90% 11-15%	50-70 15.6-18	3
	91-95% 5-10%	70.1-90 11.1-15.5	4
	>95% <5%	90.1 and above 11 and below	5

Q14	Percentile	Answers	Error Score given
	26-75%	1.91-4.1	0
	76-80% 21-25%	4.2-4.6 1.86-1.9	1
	81-85% 16-20%	4.7-4.9 1.6- 1.85	2
	86-90% 11-15%	5-5.8 1.46-1.5	3
	91-95% 5-10%	5.9-7.2 0.8-1.45	4
	>95% <5%	7.3 and above 0.7 and below	5

Q15	Percentile	Answers	Error Score given
	26-75%	162.6-167.1	0
	76-80% 21-25%	167.2-167.4 162.56-162.59	1
	81-85% 16-20%	167.5-167.58 162.51-162.55	2
	86-90% 11-15%	167.59-168.5 158.7-162.5	3
	91-95% 5-10%	168.6-170.15 157.4-158.6	4
	>95% <5%	170.16 and above 157.3 and below	5

Q16	Percentile	Answers	Error Score given
	26-75%	456.6-1360.9	0
	76-80% 21-25%	1361-1500 453.8-456.5	1
	81-85% 16-20%	1500.1-1850.9 397.6-453.7	2
	86-90% 11-15%	1851-1990.9 241-397.5	3
	91-95% 5-10%	1991-2900 113.4-240.9	4
	>95% <5%	2900.1 and above 113.3 and below	5

Q17	Percentile	Answers	Error Score given
	26-75%	151-765	0
	76-80% 21-25%	766-960 143-150	1
	81-85% 16-20%	961-970 100-142	2
	86-90% 11-15%	971-1500 95-99	3
	91-95% 5-10%	1501-2800 66-94	4
	>95% <5%	2801 and above 65 and below	5

Q18	Percentile	Answers	Error Score given
	26-75%	56-110	0
	76-80% 21-25%	111-160 40-55	1
	81-85% 16-20%	161-175 30-39	2
	86-90% 11-15%	176-179 28-29	3
	91-95% 5-10%	180-250 26-27	4
	>95% <5%	251 and above 25 and below	5

Q19	Percentile	Answers	Error Score given
	26-75%	68-270	0
	76-80% 21-25%	271-290 57-67	1
	81-85% 16-20%	291-299 47-56	2
	86-90% 11-15%	300-426.9 28-46	3
	91-95% 5-10%	427.0-520 16-27	4
	>95% <5%	521 and above 15 and below	5

Q20	Percentile	Answers	Error Score given
	26-75%	70.1-99.5	0
	76-80% 21-25%	99.6-99.7 68.8-70	1
	81-85% 16-20%	99.8-99.9 61-68.7	2
	86-90% 11-15%	100-115 60.1-60.9	3
	91-95% 5-10%	116-145 49-60	4
	>95% <5%	146 and above 48 and below	5

Q21	Percentile	Answers	Error Score given
	26-75%	95-470.9	0
	76-80% 21-25%	471- 490.9 86-94.9	1
	81-85% 16-20%	491-650 48-85.9	2
	86-90% 11-15%	651.1-820.9 25.2-47.9	3
	91-95% 5-10%	821- 989.9 10-25.1	4
	>95% <5%	990 and above 9 and below	5

Q22	Percentile	Answers	Error Score given
	26-75%	59.5-110	0
	76-80% 21-25%	110.1-113.9 58-59.4	1
	81-85% 16-20%	114-117.9 56.6-57.9	2
	86-90% 11-15%	118-134.9 45-56.5	3
	91-95% 5-10%	135-169.9 40-44	4
	>95% <5%	170 and above 40 and below	5

Q23	Percentile	Answers	Error Score given
	26-75%	2.31-5.85	0
	76-80% 21-25%	5.86-6.40 1.71-2.3	1
	81-85% 16-20%	6.41-7.50 1.56-1.70	2
	86-90% 11-15%	7.51-8.50 1.46-1.55	3
	91-95% 5-10%	8.51-14.39 0.86- 1.45	4
	>95% <5%	14.40 and above 0.85 and below	5

Q24	Percentile	Answers	Error Score given
	26-75%	41.1-95.0	0
	76-80% 21-25%	95.1-99.9 34.90-41.0	1
	81-85% 16-20%	100- 110.0 29.30-34.80	2
	86-90% 11-15%	110.1- 145.0 27.80-29.20	3
	91-95% 5-10%	145.1- 189.9 21-27.70	4
	>95% <5%	190 and above 20 and below	5

Q25	Percentile	Answers	Error Score given
	26-75%	2801-9400.9	0
	76-80% 21-25%	9401-9980.9 2251-2800.9	1
	81-85% 16-20%	9981-11900.9 1901-2250.9	2
	86-90% 11-15%	11901-14000.9 1781-1900.9	3
	91-95% 5-10%	14001- 19600.9 991-1780.9	4
	>95% <5%	19601 and above 990 and below	5

Q26	Percentile	Answers	Error Score given
	26-75%	241- 985	0
	76-80% 21-25%	986- 1400 191- 240	1
	81-85% 16-20%	1401- 1900 176- 190	2
	86-90% 11-15%	1901- 1980 101-175	3
	91-95% 5-10%	1981- 4399 89-100	4
	>95% <5%	4400 and above 88 and below	5

Q27	Percentile	Answers	Error Score given
	26-75%	19.50- 49.75	0
	76-80% 21-25%	49.76- 49.86 18.70- 19.40	1
	81-85% 16-20%	49.87- 57.00 14.5- 18.60	2
	86-90% 11-15%	57.01- 59.80 11.80- 14.40	3
	91-95% 5-10%	59.90- 75.00 9- 11.7	4
	>95% <5%	75.01 and above 8 and below	5

Q28	Percentile	Answers	Error Score given
	26-75%	7.00- 15.00	0
	76-80% 21-25%	15.01- 16.50 5.51- 6.99	1
	81-85% 16-20%	16.51- 18.50 4.61- 5.50	2
	86-90% 11-15%	18.51- 24.50 4.21- 4.60	3
	91-95% 5-10%	24.51- 28.99 2.71- 4.20	4
	>95% <5%	29.00 and above 2.70 and below	5

Q29	Percentile	Answers	Error Score given
	26-75%	986.0- 4700.0	0
	76-80% 21-25%	4700.1- 4900.9 975.0-985.9	1
	81-85% 16-20%	4901.0- 5600.9 960.1- 974.9	2
	86-90% 11-15%	5601.0- 7500.9 635.1- 960.0	3
	91-95% 5-10%	7501.0- 14000.9 301- 635.0	4
	>95% <5%	14001.0 and above 300 and below	5

Q30	Percentile	Answers	Error Score given
	26-75%	75.3- 245.0	0
	76-80% 21-25%	245.1- 275.0 57.1- 75.2	1
	81-85% 16-20%	275.1- 400.0 49.4- 57.0	2
	86-90% 11-15%	400.1- 780.9 48.3- 49.3	3
	91-95% 5-10%	781.0- 970.9 38.6- 48.2	4
	>95% <5%	971.0 and above 38.5 and below	5

Appendix 3.7

National Statistics Socio-economic Classification Volume 3

eight classes	**five classes**	**three classes**
1. Higher managerial, administrative and professional occupations	1. Higher managerial, administrative and professional occupations	1. Higher managerial, administrative and professional occupations
1.1 Large employers and higher managerial and administrative occupations		
1.2 Higher professional occupations		
2. Lower managerial, administrative and professional occupations		
3. Intermediate occupations	2. Intermediate occupations	2. Intermediate occupations
4. Small employers and own account workers	3. Small employers and own account workers	
5. Lower supervisory and technical occupations	4. Lower supervisory and technical occupations	3. Routine and manual occupations
6. Semi-routine occupations	5. Semi-routine and routine occupations	
7. Routine occupations		
8. Never worked and long-term unemployed	*Never worked and long-term unemployed	*Never worked and long-term unemployed

Retrieved from: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html

Appendix 3.8

Research Proposal

Abstract

Background: Decision-making is a part of everyday life. Damage to the frontal lobes has been known to affect decision-making capacities in humans. This can have serious implications for individuals in their ability to lead independent lives. The Cognitive Estimates Test (CET) has been used to assess the process of decision-making. Previous versions include questions that are not culture fair, with performance being strongly influenced by prior knowledge limited to certain sections of the population. **Aims:** To develop a new culture fair questionnaire and assess the reliability and validity of this measure. **Methods**: A normative range of answers from the population will be established, and a scale will be developed to define a 'bizarre response'. A One-Way ANOVA will be used to compare performance in those with brain injuries to a matched control group without brain injury on the test. As an exploratory analysis, we will compare those deemed by clinicians to have capacity to make decisions, with those who are not. **Applications**: This test may be used as a measure to assess capacity to make judgements and decisions in those with brain injuries.

Introduction

Judgement, or the ability to make considered decisions or come to sensible conclusions, is important for effective independent living. Judgement is the result of a process by which evidence is evaluated, chances of different outcomes assessed, and an action decided (Blanchette & Richards, 2010). Estimation and judgement underpin decision-making, where individuals choose something by drawing on knowledge from multiple sources, and select or avoid options that carry unfavourable outcomes (Blanchette & Richards, 2010). These abilities are central to being able to make reasoned decisions in everyday life.

Estimation and judgement are seen as executive skills. Executive dysfunction and associated impairments in estimation and judgement are seen by many to make a major contribution to neurobehavioural disability following acquired brain injury (Wood, 2001). The processes involved in estimation and judgement are considered to include the following: identifying the problem, holding and maintaining the problem in working memory, recalling relevant facts and information, iterative comparison and change to be able to estimate, and finalising the estimate. Thus, making a judgement, which we could define as "the process of forming an opinion or evaluation by discerning and comparing" (Merriam-Webster, 2013), involves the synthesis of information from multiple sources and is part of the bigger process of problem-solving which allows us to identify that a problem has occurred; take account of resources; recall problem solutions; trial potential solutions and monitor the effectiveness of the chosen approach (Evans, 2009).

Impairment of estimation and judgement can have serious implications for the individual and are important to assess. Assessment of a person's ability to estimate may be useful in determining whether someone has the capacity to make important decisions, such as managing their finances, regulating their behaviour in relationships and in terms of their future employment prospects. There have been a number of attempts to quantify judgement abilities following brain injury. The most common of these are cognitive estimate tests (CETs). In these, people are told that for a series of questions where the answer is difficult to know, estimated answers are required. Thus estimations are taken as analogues of real life judgements. There are many cognitive functions involved in estimation and therefore, many types of brain dysfunction will cause errors in these questions (Bullard et al. 2004).

Previous CETs include Shallice and Evans' (1978) CET and questions within the Behavioural Assessment of Dysexecutive Syndrome (BADS, Wilson et al., 1996). Shallice and Evans (1978) demonstrated sensitivity of CET to lesions of the frontal lobes. Both tests are, however, culture-specific, with performance being strongly influenced by prior knowledge, thus limiting populations for whom the questions are valid. For example, Shallice and Evans' test was developed in London and contained the question "how high is the Post Office Tower?", a question bound by cultural knowledge of the London skyline. The BADS (Wilson et al., 1996) includes cognitive

estimate questions that are likely to be easier for people with experience of the content of the questions – e.g. people who own a dog (for the question 'how long do most dogs live for'?), or people who go the dentist regularly (for the question 'How long does a routine dental check up take?). It is conceivable that people from areas with economic deprivation, or who belong to cultures where pet keeping, routine dental care or house maintenance services are not the norm may be disadvantaged.

Della Sala et al. (2003) developed a CET standardised on 175 healthy individuals in Italy. Their scale indicated a correct range of answers and ranges to indicate whether the response represented a bizarre response, one that is statistically deviant from the norm. The equation of bizarreness ratings and impairment of judgement is a useful development. Although this measure has many questions with good face validity and apparent culture fairness, some items appear highly culture bound, e.g. "Approximately how many coffees can a barman in a motorway restaurant make in one hour during rush hour?" Whilst there are some positive elements to this test, there remain problems with culture fairness.

It is proposed that culture fair questions should refer to the physical world, the human body, and cultural practices that have become global and familiar to most people. Questions in a culture fair test would place less reliance on recall of cultural information with more emphasis on immediately knowable aspects of their body and the physical world. It would be extremely useful to assess estimation skills as a proxy for broader problems with judgement and decision-making, as this may help identify people who will have difficulties with independent living and who may lack the capacity to make important decisions.

Aims

We aim to develop a set of culture fair questions assessing cognitive estimation and assess the reliability and validity of this measure. These questions will be piloted in a normative sample and a group of persons with brain injury, to demonstrate whether this test distinguishes between those with brain injuries and age-matched controls without brain injury. We will also explore performance in those with brain injuries who are considered to have difficulties making decisions regarding important aspects of

everyday living compared with those who are not. This will help us to see whether this test can clinically useful.

Hypotheses

Primary hypothesis. People with brain injury will give answers on the Culture Fair Assessment of Cognitive Estimates (CFACE) that are significantly different from those of healthy controls.

Secondary exploratory hypotheses. Those with brain injury deemed by clinicians as unable to make major decisions regarding their welfare, for example, managing household budgets or deciding where to live, will perform at a significantly poorer level on the test than those with capacity to make these decisions.

For participants with brain injury, performance on the CFACE will correlate significantly with ratings of executive functioning (the Dysexecutive Questionnaire) and in particular the executive cognitive sub-scale of this measure.

Plan of Investigation

Settings

Brain Injury Rehabilitation Trust at Graham Anderson House, Glasgow
Brain Injury Rehabilitation Trust at York House, York
Brain Injury Rehabilitation Trust at Daniel Yorath House, Leeds

Participants

Inclusion criteria normative sample. English speaking male and female volunteers who have given consent to participate.

Inclusion criteria patient sample. English speaking male and female volunteers, classified as having severe brain injuries and those deemed able to make an informed decision to participate in the project.

Exclusion criteria normative sample. First language other than English. Any prior history of psychiatric treatment suggesting disorders with potential neurobehavioural effects, any previous head injury requiring medical treatment, previous episode of unconsciousness, alcoholism, evidence of any neurological disorder (including stroke, seizures, tumours), chronic medical conditions which might affect neuropsychological function (such as cardiovascular disease, diabetes, hepatic disease or HIV).

Exclusion criteria patient sample. First language not English. Dysphasia.

Recruitment procedures

The normative sample will be recruited by sending information regarding the study to as many potential male and female participants as possible, via email, social networking sites and individual contact, initially.

The patient sample will be recruited from the Brain Injury Rehabilitation Trust (BIRT) at Graham Anderson House. If required, we may recruit from other UK based BIRT units, namely, York House in York, and Daniel Yorath House in Leeds. The clinical team refers into the project (i.e. potential recruits are identified at the Clinical Team Meeting), and potential participants will be given information about the study (by a member of the treating clinical team). The patient may take as long as s/he chooses to decide (within the period of the study). The patient will be reminded about the project after one week and if s/he has not yet decided whether or not to participate will only be reminded again if s/he indicates s/he would like to be reminded (which may be necessary for patients who may have some memory difficulties). If the patient indicates that s/he is interested in participating, the researcher will meet with the patient, and answer any questions about the project. If the patient is willing to participate, written consent will be obtained and testing will proceed. If the patient would like more time to consider taking part s/he will again be given a week and then reminded.

To obtain a well-matched control group, relatives of patients will be approached by the clinical team and invited to participate in the study. Relatives who are interested in participating will be contacted by the researcher who will answer any questions about the project. Relatives who consent to participating will then be tested. This testing may be done face-to-face, or by phone.

Measures

Participants with brain injury

- *Culture Fair Assessment of Cognitive Estimates (CFACE) – as developed for the study.*
- *Dysexecutive Questionnaire (DEX, Burgess, Alderman, Wilson, Evans & Emslie, 1996)*
    - This questionnaire is a part of the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al. 1996), and includes 20 items designed to sample emotional, motivational, behavioural and cognitive changes in someone with Dysexecutive Syndrome.
    - We will use the self and independent rater questionnaires with the brain injury group. A global measure of insight into post-injury deficits can be obtained by subtracting a patient's self-ratings from those of an independent rater. In addition, the level of correlation between scores on the CET and the executive cognitive subscale (Simblett and Bateman 2014) of the DEX will be examined.

Data from the following measures, routinely available within the clinical service from which patient participants will be recruited, will be collected to characterise the patient sample:

- *(Behavioural Assessment of Dysexecutive Syndrome BADS, Wilson et al. 1996)* is comprised of six subsets and is designed to assess skills and demands of everyday life and is sensitive to capacities affected by frontal lobe damage.

- *The Test of Premorbid Functioning (TOPF, Wechsler, 2009)* is a test that enables clinicians to estimate an individual's level of intellectual functioning before the onset of the brain injury or illness.
- *The Wechsler Adult Intelligence Test- fourth edition (WAIS-IV, Wechsler, 2008a)* was designed to measure intelligence in adults and is often used to assess cognitive functioning after brain injury. It is comprised of 10 core subtests and 5 supplementary subtests.

Control group (demographically matched friends and relatives of participants)

- *Developed Culture Fair Assessment of Cognitive Estimates (CFACE)*

The primary researcher will administer the above questionnaires.

Design

This study will be split into three stages. See below:

Phase 1. Existing CETs will be evaluated by the project team. From these measures, questions, which are deemed culture fair, will be included in our new measure. New questions will also be developed, and from these methods, a set of new items will be derived. To provide a check on whether the included items are culture-fair, colleagues of the research team who are familiar with different cultures/areas of the world will be invited to comment on whether they consider the items to be appropriate for their culture/context. Appropriate items are defined as ones that are understandable, to which it seems likely that an average person in their culture/context could give a reasonable estimate and for which the majority of people do not have highly specific knowledge. The most important feature of this test is that estimates are not based on detailed pre-existing knowledge of the specific focus of the question, but sufficient knowledge is available to make a reasonable estimate.

Phase 2. A normative data sample will be collected via a number of methods including via an online survey tool, by sending questionnaires via post, testing participants face-to-face and by phone. An initial invitation email will be send to potential participants giving information about the study. If they think they would like

to participate, they will be asked to send a reply email informing the researcher that they would like to take part. In order that the information given in the questionnaire is fully anonymous, the researcher will then send an email back to the participant with an anonymised patient identification number and link to the online survey tool (Qualtrics). Participants will be taken to an introductory page, which will again explain what they are being asked to do. They will then follow instructions to complete the demographic survey and CFACE. Freepost letters will be sent to participants who may not access computers. Finally, the researcher will also administer the questionnaire face-to-face, to as many people as possible recruited as a hospital and university workplace sample of convenience. This will allow comparisons of answers between participants completing the questionnaire in different ways to be analysed, helping to assess the validity of the measure in different media.

A sub-sample of 30 participants will be invited to complete the test again in order to examine test-retest reliability. This will be done at least one week and no more than one month after the first test. All participants will be asked when they first complete the test whether they would be willing to be contacted again to complete the test on a further occasion and the sample of 30 will be randomly selected from those who agree.

Cognitive estimates tests are scored in terms of amount of deviation from typical answers. So for the normative sample, the mean response and associated standard deviation will be calculated for each item. For each participant a deviation score will be calculated for each item and then totalled for the whole test. Mean, standard deviation and percentiles for total deviation scores will then be calculated. Participants' demographic information including age, gender, ethnicity, years of education, handedness, and home postcode will be gathered. The relationship between scores and demographic data will be assessed to see if any factors correlate with performance on the test.

Phase 3. The questions will then be tested on those with severe brain injuries and compared to a healthy, matched-control group. In order to obtain a more closely matched control group, relatives of participants will be invited to take part in the study. We will assess whether these questions distinguish those who professionals deem as having good capacity to make decisions and those who do not. This measure of

concurrent validity will help us to see whether this measure may be useful in a clinical setting.

Participants will also complete the DEX, a subtest of the BADS, to assess concurrent validity of the newly developed measure (whether it correlates well with this already well-validated measure which demonstrates sensitivity to everyday problems experienced by those with brain-injuries). The DEX provides a more general measure of the impact of executive functioning difficulties, and includes items relevant to making decisions in everyday life. Therefore we might expect aspects of this measure (those relevant to functions used in estimation), in particular those from the independent rater, to correlate with our test. This will increase our confidence that this new test is measuring what it was designed to measure. In addition, it will be used to characterise the sample at the time of testing.

Data analysis

Differences in the variances of scores between groups on the CFACE will be analysed. Levene's test, an inferential statistical test that assesses equality of variances between two or more groups, will be part of this process of assessing differences in variances.

A normative distribution curve will be derived for the range of answers given for each estimate question. A scoring system will be derived also, using the scores collected from the normative sample, where 0 is the mean, and scores above or below the mean are deviations from the mean. Individual performance will be measured using Z-scores for each estimate question, which can then be added together to assess overall deviation from the mean. Therefore, the higher the total score, the bigger the distance of that individual's performance from the mean (derived from the normative sample).

The DEX will be administered at the same time as the CFACE (to ensure time of testing does not interfere with results). Concurrent validity against executive cognition components of this measure, identified using principal component analysis (Simblett & Bateman, 2014) will be assessed. This will give information on whether the CFACE correlates with these factors.

The next stage will involve testing patients with brain injury against normal controls. The specificity vs. sensitivity of the test will also be explored to establish whether it distinguishes between those brain-injured patients with and without capacity to make welfare decisions. An Independent One-Way ANOVA will be carried out comparing people with brain injury with capacity to make major decisions regarding their welfare, people with brain injury without this capacity and healthy controls. Levene's test will be used to explore variances between scores. We will also explore whether any demographic factors significantly predict performance on the test using Pearson Product Moment Correlations and if so these factors will be used as covariates if there are significant group differences on any of these factors.

Test-retest reliability will be established by testing a sample of 30 participants from the normative sample. Correlation will be assessed between the two time points, with an expectation that the correlation should be at least 0.7. If reliability is at the expected level, then 30 people would be sufficient to produce a significant result. In addition any differences between scores at the two time points will be examined. Finally, the Bland and Altman (1986) statistical method of assessing agreement between clinical assessment methods will be used to determine 95% confidence limits for the difference between the two measurement points.

Justification of sample size

Normative sample. Consistent with previous publications, we aim to collect a normative sample size of at least 200. From the literature, it appears that four key factors influence performance on CETs: gender, education, intelligence/IQ and socio-economic status. Demographic information (gender, education and socio-economic status – estimated via postcode) will therefore be collected and analysed to examine the impact of these factors on performance. We aim to recruit a heterogeneous sample from the population representing a range of subgroups within each of these factors.

Patient sample. The majority of the previous literature has tested construct validity, looking at how well the CET distinguishes between those with and without brain injuries. A selection of these was evaluated. The papers assessed patients with a range of neurological issues including neurological, neurosurgical and neuropsychiatric

problems. The effect sizes calculated using Cohen's d, ranged from 0.05-2.90 with an average effect size of 1.0. For the first comparison therefore, the calculation will be powered on the basis of an effect size of 0.8, as this is deemed reasonable given previous publications. Our alpha level will be 0.05, therefore, using G-power to calculate sample size, we will aim for a minimum of 21 participants in each group.

For the second exploratory hypothesis, examining whether there is a significant difference between those with brain injuries who are deemed to have capacity to make decisions, and those who are not, it is not clear what the effect size will be as no previous studies have examined this, however, we it would need to be large for it to be clinically useful. Therefore, this part of the study will also be powered on the basis of a large effect size being present. This analysis will be underpowered if the effect size is not large, but the study will nevertheless provide useful information to power future investigations. This sample will be in part established using data from those patients in the first part of the analysis. These patients will be grouped by whether the clinical team including Psychiatrist deems them able, or not able, to make major welfare decisions (e.g. concerning need for care support, accommodation, relationships and work) in their daily lives. Based on numbers within each group at this stage, further patients will be recruited according to numbers needed to balance group numbers.

Equipment

Developed CFACE
DEX questionnaires
Demographic questionnaire
SPSS
Paper/envelopes for letters
Freepost envelops
Audio recording device

Health and Safety Issues

Researcher safety. Researcher will be working with people who have severe brain injuries. Due to the nature of these injuries, unpredictable and aggressive behaviours are possible.

Participant Safety. To someone unfamiliar with a CET, the fact that the questions have no exact answer, and may appear impossible, may cause confusion, lack of motivation or beliefs of not being able to do the test. This patient group may be particularly vulnerable to experiencing the above.

Actions to minimise risk.

Researcher Safety. Clinicians who are familiar with the client will be asked only to refer patients who do not represent a significant risk. The researcher will wear an alarm during visits. On the day of the assessment, a member of the clinical team will inform the researcher on the status of the patient.

Participant Safety. As above, a member of the clinical team will inform the researcher on the status of the patient on the day of the assessment. The procedures used in this study are similar to those administered by clinicians and are not usually associated with the onset of significant distress. To reduce distress regarding the questions, it will be made clear that the questions require a reasoned guess and that there is no exact answer.

Ensuring general safety while carrying out research. At BIRT and within other testing locations, local procedures to ensure staff and patient safety will be followed. BIRT have procedures to minimise risk to staff and these are adequate in the context of the proposed study. It will be ensured that the researcher:

1. Has a tour of the building by a member of maintenance staff to demonstrate location of emergency exits.
2. Receives training on fire procedures.
3. Demonstrates evidence of Disclosure Scotland clearance to work with vulnerable groups.
4. Has use of a personal alarm when working in clinical areas.
5. Is appraised of risk assessments pertaining to participants.

6. Is able to contact site supervisor or service manager for discussion of any concerns.

Ethical Issues

This project will be submitted to the University of Glasgow Research Ethics Committee. Although individuals with cognitive impairment impinging on capacity will be included in this study, only people considered able to consent to participate in this research project will be included, even though they may have been deemed under the Adults with Incapacity Act, not to have capacity to manage other aspects of their affairs. Capacity to make decisions is considered to be decision-specific. Deciding to take part in this research study, while not a trivial decision, is less complex than making decisions regarding personal welfare such as financial affairs, living arrangements or care needs. The nature of the research project is simple to understand and there are no major risks to participants, hence understanding the requirements and implications of participation is not complex in this case. Aspects of participation that the patient would also need to understand (e.g. confidentiality of data, ability to withdraw from the study and agreement to neuropsychological test data being obtained from clinical records) are relatively straightforward and supportable by the patient information for those with poor memory function. It seems reasonable then to expect that it is possible to recruit a sample of participants with brain injury, all of whom have the capacity to consent to participate in this research project but some of whom would lack the ability to make more complex welfare decisions.

All data will be fully anonymised and kept within locked-filing cabinets to ensure the safety and confidentiality of participants. The questionnaire will be explained fully to prevent any distress regarding the difficulty of the questions.

Participants' performance on the test, which will aim to assess an important function for capacity, will not be reported to the referring Clinical Team. This will ensure that their participation and performance can have any bearing on current or future decisions regarding their capacity.

Financial Issues

Equipment mentioned above
See Costing Form for more details

Timetable

What?	**Date to complete by**
Develop preliminary questionnaire	April 2014
Proposal submission	14th April 2014
Send/discuss to establish opinion on culture fairness	April 2014
Establish criteria and finalise questionnaire	April/May 2014
Send to ethics	August 2014
Systematic Review Outline	25th August 2014
Normative sample collection	September- December 2014
Establishment of bizarreness cut-off from normative sample (phase 2)	December/January 2015
Test patients and matched control group	February-April 2015
Establish concurrent validity using measures (DEX)	May 2015
Test-retest reliability	Re-test sample May 2015
Complete write up	June/July 2015

Practical Applications

The research would give rise to a new measure with demonstration of sensitivity to clinical condition. It might then be standardised more widely. The measure might be used in the clinical assessment of people after acquired brain injury, medico-legal settings, psychiatric settings and older adult settings. The measure might be referred to as the Glasgow Cognitive Estimates Test and be used to assess capacity and level of executive dysfunction.

References

Blanchette, I., & Richards, A. (2010). The influence of affect on higher-level cognition: A review of research on interpretation, judgment, decision making and reasoning. *Cognition and Emotion*, 24 (4), 561-595.

Bland, J. M. & Altman, D. G. (1986). 'Statistical methods for assessing agreement between two methods of clinical measurement', *The Lancet* ***i***, 307-310.

Bullard, S.E., Fein, D, Gleeson, M.K, Tischer, N, Mapou, R.L., & Kaplan, E. (2004) The Biber Cognitive Estimates Test. *Archives of Clinical Neuropsychology*, **19**, 835-846.

Burgess, P.W., Alderman, N., Evans, J.J., Emslie, H. (1996). Validity of the battery: Relationship between performance on the BADS and ratings of executive problems. In B.A. Wilson (Ed.), BADS: Behavioural assessment of the Dysexecutive syndrome manual (pp.18-19). Bury St. Edmunds, UK: Thames Valley Test Company.

Della Sala, S., MacPherson, S.E., Phillips, L.H., Sacco, L., Spinnler, H. (2003). How many camels are there in Italy? Cognitive estimates standardised on the Italian population. *Neurological Science* 24:10–15.

Evans, J.J. (2003) in Oddy, M., and Worthington, A. (Eds.) *The Rehabilitation of Executive Disorders.* Oxford: Oxford University Press.

judgment. 2013. In *Merriam-Webster.com*. Retrieved May, 8 2013, from http://www.merriam-webster.com/dictionary/judgment

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

Wechsler, D. (2009). Test of Premorbid Functioning. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008a). Manual for the Wechsler Assessment of Intelligence Scale-Fourth Ed. San-Antonio, TX: Pearson.

Wood, R. Ll., (2001). In Wood, R. and McMillan, T. (Eds) *Neurobehavioural Disability and Social Handicap Following Traumatic Brain Injury.* Hove: Psychology Press.