The Behavioural Dyscontrol Scale:

Validation of a Computerised Version in a

Non-Clinical New Zealand Population

A thesis

submitted in fulfilment

of the requirements for the degree

of

Master of Social Sciences in Psychology

at

The University of Waikato

by

Virginia Ruth Kendall

2013

Abstract

The objective of this study was to evaluate the validity of a new computerised version of the Behavioural Dyscontrol Scale (BDS) in comparison to the original manual version which research has shown to be a sensitive, reliable and valid measure of executive function (EF), and in particular of control over voluntary behaviour. A.J Luria deconstructed the complex construct of EF into Three Functional Units of working memory (Fluid Intelligence Factor), motor programming (Motor Programming Factor), and inappropriate response inhibition (Environmental Independence Factor) which he regarded to be predictive of a person's capacity to function independently and autonomously in their environment. This theoretical framework and demonstrated ecological utility is what differentiates the BDS from other traditional clinical measures of EF. The subjective scoring system has restricted the use of the BDS; the development of a valid and reliable computerised version would address this limitation generating a much greater depth and range of finite objective data. Participants were 38 tertiary students who completed a demographic questionnaire, the Hamilton Anxiety and Depression self-report Scale (HADS), the Integrated Visual and Auditory Continuous Performance Test (IVACPT), Trail Making Test A and B, the manual and computerised versions of the Behavioural Dyscontrol Scale. Findings showed good levels of internal reliability and construct validity for the CBDS which yielded high sensitivity and specificity across all Three Functional Units, together with a high level of correspondence to scores generated by the manual version and by the Trails and IVACPT measures. Potential clinical applications, limitations and future directions are discussed.

Acknowledgements

I would like to thank my supervisor Nicola and Librarian Jenny for their calm and consistently optimistic approach in supporting me through this research, despite a regular flow of interruptions and unexpected life events being part of the whole process. In particular I have learned that sometimes it is of more value to stop, step back and take time to look at what is happening from a new perspective rather keeping on going when you are too immersed in the process to see clearly! A useful lesson in life too! I would also like to say thank you to my two grown-up children who have provided many tasty meals and welcome distractions along the way, to my mum for her encouraging emails from the other side of the world, and to my good friends at university. The process of doing this work has been fascinating and challenging, and my hope is that it may also be of some use and benefit to others at some point in the field of neuropsychological rehabilitation.

Contents

List of Tables

Introduction

Executive Function

Executive Function (EF) as a neuropsychological construct broadly describes higher level cognitive processes that control capacity for behavioural self-regulation and the execution of mental and motor actions (Giovanello & Vaughn, 2010; Eastvold, Suchy, Wilson, Whittaker, & Strassberg, 2007). EF has been described as a “most subtle and central realm of human activity” and is associated with the ability adaptively respond to novel situations in an appropriate, socially responsible and effective self-serving manner (Lezak, p. 612, 2004).Garcia-Barrera, Kamphaus, and Bandalos, (2010) comment that there is no equivalent lay concept which accurately captures the sophistication and complexity of this construct, and that consequently among researchers there are many definitions. In essence executive functions are generally thought to be unique to humans because they involve such mental activities as suppression of reflexes, programming and performing complex series of movements, mental flexibility (for example set shifting and resisting interference), the ability to reverse automatised series, fluency, self-correction, creativity, and personality features such as initiative, self-monitoring, and insight (Eckland-Johnson, Miller, & Sweet 2004). It is recognized that the “supra–ordinate function” of this area of the brain is structuring goal directed behaviour and the capacity to compare the result of an action with its original intention. This capacity to regulate and monitor behaviour is central to executive function. For that reason impaired executive processes can have a profound impact on a person’s ability to function in the world on many different levels and in many different contexts (Spitz, Ponsford, Rudziki, & Mallerl, 2012).

Within the field of neuropsychology EF is increasingly being viewed and studied as a unique multicomponent construct comprising separable yet correlated processes identified as working memory, appropriate response selection and self-regulation. Research has found these processes to have high levels of convergent and discriminant validity and be linked to frontal lobe brain structures (Eckland-Johnson, et al., 2004; Giovanello & Vaughn, 2010). The construct emerged out of research looking at people with frontal lobe damage who exhibited deficits in capacity to plan, organise and implement functional goal-oriented behaviour (Giovanello &, Vaughn 2010). The anterior area of the frontal lobes is the pre-frontal region which can be broadly sub-divided into the orbito-frontal cortex and the dorso-lateral cortex; these are thought to be responsible for the “maintenance and execution of abstract thought, reasoning and inhibition of responses” (Martin, 2006, p. 166). Frais, Dixon and Strauss (2006) refer to neuro-imaging studies showing “shared neural networks activated by multiple tasks that share the same underlying processes” (p.212). The authors concluded that their research indicated the existence of a specific pre-frontal network engaged in a diverse range of cognitive demands including response selection, stimulus recognition, working memory, and problem solving.

Blair, Greenberg, Willoughby and Wirth (2012), define EF “as a supervisory system that is important for reasoning ability, and the integration of thought and action” (p.226). Disruptions to the pre-frontal cortex from trauma or illness can therefore interfere with a person's ability to plan and guide their own behaviour (Grigsby & Kaye, 1996). According to (Blair et al., 2012) the construct of EF contains sub-categories of specific interconnected information processing mechanisms that coordinate and enable the synthesis and resolution of conflicting information. These are described by Lezak (2004) as volition, planning, purposive action and effective performance. Blair and others (2012) refer to these sub categories as working memory, inhibitory control, and attention shifting. Working memory can be defined as the ability to hold in mind some information whilst simultaneously updating it and utilising it in some way. Inhibitory control can be defined as the inhibition of atomised responding when engaged in a task. Attention shifting can be defined as the ability to shift cognitive set between distinct yet related domains of a given task. Pronk, Karremans & Wigboldus (2011) noted that despite the differences between these aspects of EF control, shared fundamental underlying processes were in operation. "Specifically, they all involve the capacity to focus attention on relevant information and processes while inhibiting irrelevant ones" (p. 828).Volition refers to the capacity to initiate appropriate responses to external stimuli. Deficiencies in this capacity may result because of interruption to cognitive and affective processes. Planning refers to the ability to conceptualise changes from the current situation and hold sustained attention whilst utilising memory functions and impulse control. Purposive action refers to the capacity to translate verbalised intention into effective performance. Motivation, knowledge, or capacity to actually perform an activity is not sufficient to ensure the carrying out of the intention if the programming functioning is impaired. Self-regulation refers to the ability to convert intention into action together with flexibility and the capacity to shift perceptual organization to the changing needs of the situation (Spitz et al., 2012; Lezak, 2004).

Effects of impaired Executive functioning

The location of the frontal lobes, in the frontal upper area above the lateral sulcus and in front of the central sulcus renders this region particularly susceptible to injury (Suchy, Eastvold, Wilson, Whittaker & Strassberg, 2007). The frontal lobes are anatomically highly complex and according to Martin (2006) are recognised as the most recently developed area of the cerebral-cortex. They have a protracted development and are known not to reach maturation until early adulthood (Taylor, Barker, Heavey, McHale, 2012). Consequently behavioural inhibition, impulse control, attentional switching, planning and perspective taking are attributes likely to be deficient in mid to late adolescence. Immaturity of the pre-frontal cortex in combination with hormonal development, sexual maturation, and dynamic social, intellectual and emotional change means this age group has a significantly increased likelihood of engaging in injurious and risk taking behaviour. Substance abuse and Traumatic Head Injury (TBI) can both interrupt this period of neurological maturation precipitating structural brain changes that negatively impact the developmental processes of neuronal rewiring and synaptic pruning. These processes are thought to be necessary in achieving optimum functional connectivity between brain regions by early adulthood (Taylor et al., 2012).

Research is increasingly recognising the significant impact impaired executive processing can have in reducing a person's capacity to function successfully in the context of daily living. Elkind, Rubin, Rosenthal, Skoff and Prather (2001) state that deficits in executive functions comprise the following areas of dysfunction: Mental inflexibility, difficulty in shifting attention from one concept to another and processing previously acquired information. EF dysfunction can arise as the result of a wide range of conditions for example, TBI, dementia, the ageing process, Alzheimer's disease, cardio-vascular accidents (CVA), tumours, Parkinson's disease, Wilson's disease, Huntingdon's disease, motor neurone damage, psychosis, schizophrenia, alcoholism, mood disorders, and epilepsy (Martin, 2006). In (2012) Spitz and others asserted that TBI may result in profound changes to a person's social, behavioural, emotional and cognitive functioning due to problematic alterations in memory, information processing, and executive functions. As mentioned earlier the prefrontal cortex is an area of the brain which continues to develop into early adulthood and tends to diminish in capacity into older adulthood, therefore deficits must be evaluated within a developmental and environmental context (Blair et al.,, 2012). Research looking at children, showed strong associations between executive function, academic achievement and learning related social skills which suggests that EF can be an important predictor of school readiness and potential behavioural difficulties within the school environment (Blair et al.,, 2012).

Lezak (2004) refers to *dysexecutive syndrome* occurring as a result of failure in the *supervisory attentional system,* described as a central attentional controller that “selects and operates strategies for maintaining and switching attention as needs arise” (p.615). Following TBI it has been hypothesised that poor response to behaviour modification treatment may be due to behavioural dyscontrol caused by the presence of severe *dysexecutive syndrome* (Alderman, 1996). Individuals with diminished capacity to initiate appropriate goal directed behaviour in response to external stimuli might not be responsive to traditional operant learning procedures used as compensatory treatment interventions. Difficulties with implementing goal oriented, purposeful actions and inhibiting inappropriate responses has implications for the effectiveness of cognitive behavioural interventions during treatment because the individual is likely to have greater difficulty in learning new behaviours and using self-monitoring (Spitz et al.,, 2012).This can impact on their rehabilitation and ability to function adaptively creating difficulties associated with returning to previous employment, education or pre-morbid relationships and family life (Spitz et al.,, 2012).

Individuals with diminished executive capacity may be unresponsive to verbal and non-verbal cues and at times demonstrate inappropriate behaviours that may negatively influence interpersonal relationships. Spitz et al., (2012) noted that challenges related to reintegration into the community are often experienced due to poor self-regulation and behavioural changes. Lack of insight and self-awareness (which may be a newly developed area of weakness for the individual) means that they are also less equipped to implement strategies that may help compensate for these deficits. Lezak (2004) noted that novel behaviours are more susceptible to the effects of impaired planning capacity than routine activities, because the planning function is not as necessary when performing proceduralised tasks that are over- learned and automatic. The impact on everyday living of impaired EF may easily go unrecognised because this area of brain functioning is complex, hard to define, and poorly understood by the majority of the lay population.

The diverse functioning of the frontal lobes means damage can cause a wide range of impairments. People with executive deficits may experience a range of the following difficulties: Lack of insight in how to integrate or understand feedback from others; Difficulty in modifying response patterns; Impulsivity and difficulty modulating ingoing behaviour; Difficulties in organizing thoughts and executing intentional goal directed behaviour; *Pathological inertia* is where a person describes an intended behaviour but never acts it out; Difficulties with formulating generalisations among specific events or principles; Impairment in abstract thinking; A concrete thinking style, literal mindedness and loss of perspective (Barrera et al., 2011; Lezak, 2004). Perseveration (defined as a persistent behaviour that continues even when it becomes interpersonally or circumstantially inappropriate) errors are common in those with executive deficits. They are the resulting responses of inflexible non-adaptive actions. According to Lezak (2004) perseveration is indicative of difficulties with self-regulation arising from an inability to shift thought (set shifting) and behaviour to the changing needs of the moment. This results in an involuntary repetition of ideas or experiences without the presence of appropriate stimuli (Lezak, 2004).

Chan (2001) suggests that the effects of executive dysfunction may be evidenced at a cognitive behavioural level when performing everyday tasks as well as a cognitive level as evidenced by traditional assessment measures. Chan goes onto comment that attentional lapses, absentmindedness, and mistake making are all aspects of attentional control and action disorganization experienced by the normal population. Whilst the degree of deficit may not be at the level of impairment, from a functional perspective it may be viewed as a base-rate of symptoms comparable to those of individuals with frontal lobe damage. Everyday routine activities although requiring a relatively low level of demand, do stand somewhere along the continuum of executive demands (Chan, 2001). Chan (2001) reported that individuals with frontal lobe damage performed within the normal range on neuropsychological tests of executive function, yet often the same individual consistently reported experiencing difficulty in performing everyday life tasks (Chan, 2001). This study reflects the inconsistency between results of actuarial assessment and observed behaviour often found in people with impaired frontal lobe functioning. Obtaining useful, accurate, and ecologically valid data is therefore an area of particular complexity and challenge when developing and administering measures.

Further to this, Elkind et al., (2001) describe problems with Instrumental Activities of Daily Living (IADL's) such as inadequate self-care skills (for example personal grooming and hygiene), initiating and completing of goal directed actions (for example taking regular medication), and difficulties with concentration. Research by Stilley et al., (2010) has shown that adherence to medication regimes by those with chronic health problems like diabetes, heart disease, hypertension and breast cancer, can be predicted by poor attentional capacity and working memory. The authors suggest that targeting cognitive functioning can help to identify those individuals who may be at high risk of non-adherence (Stilley, 2010).

Giovanello and Vaughn (2010) investigated the relationship between three underlying executive processes (working memory, task switching and inhibition) and two aspects of IADL'S, performance and self-report in a sample of older adults. The aim was to specifically assess the influence of EF in daily life. As expected they found a strong and significant relationship between executive processes and performance based IADLs as compared to self-report and executive processes. In the older population a positive relationship is recognised between levels of global cognition and IADLs, however the results of this study demonstrate that EF may be a more reliable predictor of functional outcome (Giovanello & Vaughn, 2010). This demonstrates the ecological utility of specific measures to predict an individual's daily functioning using EF assessment.

Assessment of Executive Function

The complexity and multi-causality of impaired executive functioning is reflected in the variety of measures used for assessment. Neuropsychological testing aims to examine a range of cognitive abilities, including memory, speed of information processing, language, attention, and executive functioning. Findings from assessments help facilitate appropriate intervention, measure treatment efficacy, and guide clinical judgment.

Accuracy and reliability of assessment is therefore essential and carries serious implications for diagnosis, treatment and outcome. Many traditional actuarial cognitive assessment techniques have been found to inadequately reflect observed real life behaviour in those with impaired frontal lobe functioning (Chan, 2001; Cope, 2005; Spitz et al., 2012). Objective, quantifiable, and ecologically valid data has historically proved difficult to obtain, however as understanding around the complex system of interrelated cognitive processes increases, so does the importance attributed to the work of A.R Luria in shaping the current conceptualisation of Executive function.

Discrete versus Integrated Systems

From a historical perspective measures of neuropsychological assessment have changed considerably in response to evolving theories, social construction and empirical research. The development of modern tests of EF has been significantly informed by the work of A.R. Luria a Russian neuropsychologist influential from the early to late twentieth century. Essentially Luria proposed that the brain operates as a whole co-ordinated functioning system. Therefore, if any specific area is dysfunctional this will impair the overall functioning of the brain making it likely that a range of behaviour will be disrupted (Plaisted et al., 1983). Luria formulated and developed the psycho physiological theory of cognitive functioning between 1922-1980 (Glozman, 2007). Russian neuropsychology was influential from the eighteenth century onwards and according to Glozman (2007) focused on many areas fundamental to modern neuropsychology (such as aphasia, apraxia, and agnosia). Lebedinsky pioneered the principle of integration and differentiation of separate brain areas as a simultaneous function. Vasilenko developed the methodological approaches of using the patient's own activity, initiative and residual compensatory mechanisms to examine the complex processes involved in aphasia. This research was novel in that it focused on factors affecting the rehabilitation potential of people with aphasia. It formed the background to Luria's work which began around 1920, beginning the advent of what is now regarded as the 'Lurian period' in neuropsychological research history (Glozman, 2007).

Luria observed that the pre-frontal cortex and anterior cingulate were particularly active when new problems were being solved utilising decisive higher forms of attention and vigilance. Once tasks became proceduralised these areas of the brain were observed to be less active (Lezak, 2004). The origins of labels used today to describe EF such as volition, purposive action, and planning can be clearly seen in Luria's theory that proposed the brain could be divided into Three Functional Units for the purposes of "any type of mental activity" (Languis & Miller 1992, p.494). These Three Functional Units were: Working Memory, Appropriate Response Selection, and Inappropriate Response Inhibition. Luria was primarily interested in explaining the mechanisms that linked elementary and higher forms of psychological organization and processing in healthy adults. Luria attempted to identify the mechanisms of cognitive deficiency that appeared in different types of early abnormal ontogenic development and different forms of brain injury (Glozman, 2007; Plaisted, Wilkening, Gustavson & Golden, 1983).

Pluripotentiality is a key component to this theory. Luria proposed that separate areas do have discrete functions differentiating them from other areas and do not contribute equally to all behaviours, however all areas are interrelated. Therefore, if any specific area is dysfunctional this will impair the overall functioning of the brain and it is likely that a range of behaviour will be disrupted (Plaisted et al.,; 1983). Whilst different areas of the brain may have specific functions each are incapable of separate operation, and all behaviour is actually produced by the complex interaction of numerous separate parts. This multifunctionality of independent parts is significant in the event of trauma to one area because overall functioning can often be maintained using an alternative functioning system (Glozman, 2007). This also means that trauma to one area of the brain can affect functioning in that specific region or equally have more systemic consequences.

The existence of these separate functioning yet interdependent areas of the brain is relevant to the determination of accurate neuropsychological assessment. It means that deficits are best measured linked to a primary deficit, rather than in isolation and the systemic consequences and compensatory alternative functioning mechanisms considered (Glozman, 2007). According to Glozman (2007), Luria also emphasized the importance of qualitative observation and value of identifying preserved strengths in his theory of neuropsychological rehabilitation. The novel concept of using this information to predict an individual’s potential for successful reintegration into society was created and has increasingly informed modern research approaches to neuropsychological assessment, particularly in the area of predicting functional outcome following traumatic brain injury (Spitz et al, 2012). Luria's approach has since been developed by subsequent generations of neuropsychologists around the world.

Measures used to assess Executive Function

Luria's work is embedded within modern approaches to the assessment of Executive Functioning. He highlighted the importance of ecological validity of testing, and as awareness of the significance of this construct has increased, so has effort invested into developing improved measures of assessment. The field of EF testing is continually developing in response to the increasingly recognised limitations associated with traditional measures. The trend of modern assessment is in line with Luria's work to focus more on predicting reliable functional outcomes than discrete measures of performance that may lack utility and relevance to real life.

The accurate, discrete measurement of complex executive processes and the production of reliable and useful data remains a challenge to the field of neuropsychology. "However, the importance of improving the characterisation of cognitive impairment is essential given its strong association with functional outcome" (Spitz et al., 2012, p 604). The potential disparity between real-life functioning of a person and recommendations made following formal assessment of executive functioning by clinicians is discussed by Elkin et al., (2001). The authors highlight the lack of real life interferences, stresses and multiple demands in traditional neuropsychological testing settings which can bias results "The artificiality of executive functioning assessments limits their capacity to predict real life functioning" (Elkin et al., 2001, p. 491). Alderman

(1996) reported test measures of general intelligence, memory and frontal lobe functioning did not effectively discriminate between individuals who displayed symptoms of *dysexecutive syndrome* and those who did not.

Due to the complexity of the construct, EF can be assessed in a multitude of ways and dysfunction can occur at any stage of the behavioural sequence. Disruption can take place at any stage; volition, planning, purposive action, or effective performance (Lezak, 2004). It is unusual for only one of these processes to be affected due to the 'orchestral nature' of this area of the brain; according to Lezak (2004) "defective executive behaviour typically involves a cluster of deficiencies of which one or two may be especially prominent" (p.611). Designing measures aimed at being sensitive to these 'prominent' deficiencies has become the increasing focus of modern research, and computerised assessment is one area of this development.

Computerised Testing – limitations and advantages

Computerised neurocognitive testing has increased in use over the past decade, particularly with healthy children, adults and elderly populations (Iverson, Brooks, Ashton, Johnson, & Gualtieri, 2009). Advancement in computerised technology in recent years has led to an increase in the development of computerised cognitive test batteries. Vispoel (2000) cited guidelines for evaluating the inter-changeability of computerised and conventional psychometric testing specified by the American Psychological Association published guidelines (1986). These were that similarity of scores across modes of administration and rank ordering of scores would suggest the manual and computerised versions were fundamentally measuring the same traits. However some research (Cope et al., 2005) suggests that equivalence should not be expected and is not possible due to the inherent differences in the two modes of administration. Cope and others (2005) concluded both modes of testing have aspects which cannot be fully and accurately replicated by the other and that this should be recognised.

A salient advantage of computerised neuropsychological testing is standardisation, accuracy and consistency of information recording and presentation of data. Relatively large amounts of data can be saved directly onto disc with high test security and space efficient information storage. Costs of administrating a battery of tests can be considerably less than for traditionally administered versions, with training administrator costs and reduction in administrator bias also being reduced. Possibilities for using different languages to present tests provide wide ranging potential cross-cultural advantages for computer administration. Accurate measurements of time sensitive tasks using small time units such as milliseconds increases capacity to measure areas of performance not possible through conventional modes of administration, such as latency, strength and variability in response patterns, producing data immediately for interpretation. High face value and performance motivating effects exist for clients who enjoy using computers or feel uncomfortable with the interpersonal aspect of traditionally administered tests. However negative self-evaluation and anxiety can occur in participants unfamiliar with computers. There is a limit to the types of material that can be presented via this medium, for example assessing qualitative information such as behaviour, mood, frustration tolerance and insight tends to be restricted by the use of multi-choice or yes/no answer questions. Technical inconsistencies in quality of visual graphics and sound quality can occur and the cost of designing measures, maintaining the software and developing programs for analysis of data can be considerable in comparison to traditionally administered tests. Potential for practice effects exists and to date computer tests have primarily only been used with individual participants.

Paul, Lawrence, Williams, Richard, Cooper, & Gordon (2005) discuss the relative merits of computerised and manual test batteries. They examined the validity of the Integ-Neuro TM a newly developed measure of cognitive functioning in 50 healthy adults. Highly significant correlations between the two modes of administration were found. The authors concluded the computerised battery has the potential capacity to detect clinically meaningful declines in cognitive function associated with degenerative disease and brain injury (Paul et al., 2005). Among the advantageous features identified were the standardized instructions which utilised both visual and auditory cues, practice trials before test trials, and semi-automated scoring procedures. Test retest reliability was found acceptable across all measures. The issue of computer familiarity was identified as a potential confounding variable, with this being influenced by level of education and age. Paul et al., (2005) noted that the high level of education within the sample of participants shed some doubt over the generalisability of findings to individuals with lower education levels. Paul et al., (2005) go on to suggest that future research using less highly educated participants would be useful in order to determine the usefulness of the Integ-Neuro battery in clinical situations.

In a comparison of score compatibility and respondent preference between computerised and paper-and-pencil versions of the Self- Description Questionnaire, Vispoel (2000) found results supported comparability of the scores. In relation to this he commented that "psychometric evidence alone, however, may not be a compelling enough reason to use the computerised SDQ-111 if respondents had unfavourable attitudes about taking it" (p.140.) This comment alludes to the potential impact of participant preference for test mode on test performance. Vispoel (2000) asserted that the majority of respondents strongly favoured the computerised version over the pencil-and-paper version. This preference was indicated on the basis of less fatigue being associated with the computerised version. The validity and reliability of the two measures were both perceived as equal by respondents. Keith et al., (1998) reported the computerised administration had a high level of acceptability with participants. Weber, Fritze, Schneider, Kuhner, and Maurer (2002) viewed acceptability of the computerised administration as a significant motivational factor in relation to validity and reliability Kobak, Reynolds, and Griest (1994) found that in a sample of 121 participants with and without affective disorders, those with anxiety or depression preferred the clinician administered measure as compared to those without anxiety or depression who indicated no preference. Previous experience with computers, attitude toward computers and educational level were found to effect patient-computer interaction (Weber et al., 2002). A significant relation was found between computer attitude and "relaxation during computerised assessment", in particular biased findings in measures of attention were found especially in participants with depressive disorders (p.128). Bandura's theory of self-efficacy relates to the participant's belief that they can or cannot complete a computerised task. The strength of this belief is hypothesised by (Albert, Browndyke, Malone, Schatz, and Cohen & Gouvier, 2002) to play a fundamental role in determining the degree of "computer specific negative affect that a person may experience" (p.210). Computer related anxiety was found to be significantly related to higher error scores and longer response times, particularly in individuals with some history of impairment (Albert et al., 2002) .Chervinskaya (2005) stated that elevated levels of anxiety confounded results on the computer administered battery. Elkind et al., (2001) comment that the examiner presenting instructions may allow the participant to adaptively compensate for executive difficulties in manually administered measures which may bias results.

Familiarity with computers and influence on neurocognitive test performance was examined by Iverson, Brooks, Ashton, Johnson and Gualtieri (2009). A significant effect was found on the Stroop Test and Shifting Attention Test between those reporting "some" computer use and those reporting "frequent" computer use. Significant differences between the two groups were also reported on Psychomotor Speed, Reaction Time, Cognitive Flexibility and Cognitive Attention, with the "frequent" use group attaining better performances.

Roebuck-Spencer, Reeves, Blieberg, Schwab, Salazar, Harvey, Brown and Warden (2008) examined the influence of demographic factors on computerised testing performance in a military population finding age and gender to be important factors to consider when using reaction time based computerised assessment. Reaction time increased with age, and woman focused more on accuracy than speed, with men prioritizing speed over accuracy.

Chervinskaya and Shchelkova (2005) point out that the methodology required to achieve equivalence between pen and paper and computerised assessment needs to be individualised for each measure. They concluded specialised technical knowledge in collaboration with expert clinical knowledge is required to standardise and formalise methodology and interpretation.

Buxbaum, Dawson, and Linsley (2012) conducted research into the validity of the Virtual Reality Lateralized Attention Test (VRLAT). This was an easy to administer computerised measure assessing hemi-spatial neglect in right hemisphere stroke patients. The VRLAT was found to demonstrate strong specificity and sensitivity, with minimal practice effects, providing ecologically valid data that outperformed traditional pen and paper tasks in the prediction of real world functional outcome. The authors concluded the VRLAT to be a sensitive, reliable, and valid measure that had the advantages of being easy to administer, not requiring any specialised equipment, and valuable as a clinical and research psychometric tool (Buxbaum et al., 2012).

The BDS-EV was used to examine the relationship between motor programming (planning, learning and control) and executive functioning (Eastvold et al.,2007). This research found that only the motor planning component of motor programming was related to executive function.

Moreover, motor learning and control were subject to the effects of task complexity and novelty whereas motor planning was not. The author's suggest that the relationship between EF and motor planning ability can be conceptualized from an evolutionary perspective. Advance planning and inhibition of immediate motor responding are specifically associated with higher order species and survival (Eastvold, 2007). The separation between some of these fine levels of executive processing should be far more possible using computerised technology than manual assessment and subjective scoring systems.

In summary many factors could potentially influence the equivalence of computerised assessment measures with clinician administered assessment. These include participant's experience with computers, anxiety, negative affect, attitude to computerised testing, preference for personal interaction with the administrator, age, education and gender. Other potential confounding factors include quality of software and test design, mode of data extraction, quality of analysis and interpretation. It is important that test scores accurately reflect what they are intended to measure, and effects of unwanted variables minimised. Different formats may be more or less useful in different clinical situations and with different populations and age groups (Blair et al., 2012). It seems clear that environmental and personal biases that participants may bring to the assessment are not overlooked. Appropriateness of computerised testing depends on the population being tested, purpose of the test, the type of measure being administered (questionnaire or neuropsychological) and the context in which it is being used.

Traditional Measures of EF

Given the multi-functionality and interrelatedness of Executive function, there is no one stand-alone gold standard measure for this construct, rather a multitude of different measures have been developed to assess different prominent processes. There are a range of test batteries and also specific tests designed to measure specific aspects of EF.

As previously mentioned, one criticism of traditional measures assessments has been a failure to differentiate between the capacity for self-regulation and the transferral of verbalised intention into successful performance, which means people sometimes find real world coping harder than may be indicated by the assessment measure. Commonly administered tests to identify deficits in EF include the Wisconsin Card Sorting Test, the Stroop Colour and Word Interference Test, the Weshler Freedom from Distractibility factor, Trail Making Test, the Towers of Hanoi and London, and the Delis-Kaplan EF System (DKEFS) which is a battery of nine tests each developed as stand-alone measure; however research has shown that subtle weaknesses are often not reflected in results from these measures based on early neuropsychological models (Barrera et al., 2011; Chan, 2001). Criticisms of some of these tests included lengthy administration time, requiring apparatus, and limited practicality which are important for widespread usability.

Due to the complexity of many EF processes, deficits can remain undetected in individuals attaining high scores in other areas of cognitive testing, for example intelligence and memory. Deficient performance can be misattributed to behavioural causes, potentially resulting in inappropriate clinical interventions (Groth-Marnat, 2003; Blair et al., 2012). The importance of ecological validity and developmental stage appropriateness in EF testing is being increasingly acknowledged and represented in modern assessment design. According to Barrera et al., (2011) traditional global measures have neglected to account for these “everyday behavioural components of EF” (p.3) which Luria recognised a long time ago. As mentioned before, a more recent ecological psychometric approach has been to focus on the “analysis of the everyday behavioural components of EF” (Barrera et al., 2011, p.3). Barrera (2011) cites The Frontal Systems Behaviour Scales and the Dysexecutive Questionnaire Behaviour Assessment as two examples from the adult literature, and The Behaviour Rating Inventory of EF (BRIEF) from the children’s literature. In relation to this the author comments on the scarcity of options for the behavioural assessment of children’s EF. Blair and others (2012) also refer to the lack of specific and sensitive measures for children that account for the protracted development of the prefrontal cortex. Blair and others (2012) discusses the need for the development of a computerised battery of EF tests for children as this age group is characterised by rapidly changing pre-frontal neural network organisation, and poor executive functioning is associated with academic and behavioural difficulties in childhood. Performance on complex tasks with high processing demands has been shown to increase throughout adolescence not maturing until early adulthood (Blair et al., 2012).

Decline in pre-frontal functioning occurs with the ageing process. Perseveration is one of the “hallmarks of impaired capacity to shift responses easily and appropriately" (Lezak, 2004, p.632.). One of the most common types of perseverative response is the inability to terminate an elementary movement. It has been noted that perseverative responses often do not emerge until the latter half of tasks, thus it is important when administering tasks to continue doing so for a long enough period of time for inappropriate responding to emerge. Disassociation between being able to verbally repeat the instructions and implement the instructions correctly is often found when administering these tasks (Lezak, 2004).

The Frontal Assessment Battery (FAB) developed in 2000 by Dubois, Slachevsky, et al., comprises a series of simple tasks taking approximately ten minutes to administer. According to Lezak (2004) the FAB has shown good inter-rator reliability and internal consistency. This measure contained elements of Luria's model using a motor sequencing task (fist - edge- palm), a sensitivity to interference task, item generation and an environmental autonomy tasks (highlighting task initiation behaviour). Grooming and interpersonal style can be indicative of social awareness capacities (Lezak, 2004), and qualitative observation during the neuropsychological examination is suggested as the best method for assessing insight. Awareness of errors, participant's comments around their own performance level, attitude to mistakes and compensatory efforts can all be observed by the clinician (Lezak, 2004; Grigsby & Kaye 1996). Cognitive flexibility has been shown to have a relationship to empathy in persons with brain lesions; Luria linked insight to the working memory dimension of EF (Lezak, 2004).

Lezak (2004) describes the Tinker toy Test (TTT) (Lezak, 1982) as a reliable and valid measure of purposive action. The TTT facilitates the initiation, planning and implementing of a relatively complex construction activity. Lezak (2004) comments that these functions “typically remain unexamined, although they are absolutely essential to the maintenance of social independence in a complex society” (p.621). Lezak (2004) comments on the paradoxical nature of assessing executive deficit that necessitates the need to structure a situation measuring the extent to which individuals can make structure themselves. Systematic assessment of the different aspects of EF needs to be conducted looking at the stages at which a breakdown in behaviour occurs. This theoretical approach is supported by Diesfieldt (2004) who stated that executive functioning is a “multidimensional construct” and therefore multiple indicators need to be applied in order to accurately measure functioning (p.1065).

The Behavioural Dyscontrol Scale (BDS) is one test developed to assess the specific and separate dimensions of the EF construct using rapid finger sequencing (as if piano playing), hand sequencing (fist, edge, palm), and tasks requiring the participant to make converse responses to that of the examiner which can all elicit impaired motor regulation and inability to self-correct (Lezak, 2004). The literature strongly suggests there is a lack of measures specifically designed to quantify and separate these higher level multidimensional aspects of behaviour. The Behavioural Dyscontrol Scale is one test that has been developed informed by Luria’s theory of pluriplasticity and it is it a well validated predictor of functional outcome following TBI.

The ability to discriminate between the capacity for self-regulation and the transferral of verbalised intention into successful performance is a key aspect of the BDS differentiating it from other measures of executive function.

The Behavioural Dyscontrol Scale (BDS)

Manual Version

The BDS is a brief, standardised and easy to administer measure designed specifically to assess the more subtle and fine grain level components of neurological functioning considered by Luria to be so important. Traditionally, executive test measures have tended to focus on memory and language using written and verbal output, whereas the Behavioural Dyscontrol Scale (BDS) is specifically intended to examine the regulation of purposeful behaviour. The BDS is based on the work of Luria and adapted from his theory of frontal lobe functioning (Cope et al., 2005; Grigsby, Kaye, & Robbins, 1992). It was originally designed to assess geriatric capacity for functional autonomy, and specifically intended to examine the regulation of purposeful behaviour. Research within this population has shown the BDS to have high rates of reliability, internal consistency, and validity showing the measure to be a strong predictor of functional independence. It has also been found to highly correlate with the Trail Making B which is an established and well validated measure of EF, in particular working memory (Cope et al., 2005) as well as demonstrating high inter-rater reliability and good test-retest reliability in a non-clinical elderly population.

The measure directly assesses behavioural control, is relatively independent of disorders of affect and cognition, is brief and easy to administer. It aims to predict how well an individual will be able to carry out the activities of daily living (ADL's) which can assist in planning individualised and appropriate rehabilitative treatment in a variety of contexts (Grigsby et al., 1996). The ecological validity of the BDS gives it a unique point of difference from other traditional measures of EF.

According to Stuss et al., 2002 (as cited by Johnson, 2004), research shows that executive functions can be usefully "fractionated into separate abilities" (p. 395). Discerning between these separate abilities has been increasingly recognised as a major limitation associated with traditional EF tests. Hall and Harvey (2008) describe the BDS as being differentiated from other measures of EF by its capacity to recognise and asses a person's control over voluntary behaviour and disparity between verbalised intention to complete a given task, and the effective implementation of the intended response. In this way it aims to discriminate between individuals who due to neural dysfunction are unable to regulate their behaviour from those who have the neural capacity to regulate their behaviour, but do not for some other reason. Hall and Harvey (2008) used the BDS to examine possible differences between samples of patients with Alzheimer's and Vascular Dementia in cognitive functioning. It was reported that performance on problem-solving tasks requiring the voluntary control of behaviour could help to differentiate between mild cognitive impairment and normal ageing. Myers, Grigsby, Teel, and Kramer, (2009) examined the addition of the BDS to nursing assessment of functional autonomy capacity of patients admitted to a rehabilitation unit following fractures, medical and surgical conditions. The study concluded that using the BDS to indicate level of EF enhanced the predictive accuracy of the existing model being use, and suggested " the inclusion of this measurement of executive cognitive function may significantly enhance the accuracy of nursing prognoses " (p.264).

The BDS is comprised of nine tasks altogether, categorised into three factors based on Luria's Three Functional Units. Four motor learning and motor programming tasks are designed to detect volition and sustained attention grouped under the Motor Programming Factor (MPF). This index is designed to target the ability to purposely plan and implement a goal directed action in the absence of external cues. Two go/no-go tasks designed to tap into impulsivity and environmental dependency grouped under Environmental Independence Factor (EIF). This index aims to assess the capacity to inhibit inappropriate responses to changing environmental stimuli. The remaining three items measure working memory, reason, insight, and ability to adapt responding to feedback and are grouped as Fluid Intelligence Factor (FIF). This index is aimed at identifying ability to hold and manipulate information whilst resisting interference and adapting to novel situations (Eastvold et al., 2007).

Each task is subjectively scored out of a maximum of three points and a perfect performance receives twenty seven points. Detailed information on scoring and administration is provided in the manual (Grigsby & Kaye, 1996) Factor analysis showed that of these three functional areas FIF and MPF were the most “robust factors” (Suchy et al., 2003). Normative data was based on 1310 adults with a mean age of 74. Age related decline in scores after age 60 were shown. A study from 43 young adults showed a mean score of 17.6 (Grigsby, Kaye & Robbins, 1995). It should be noted that the individuals age, physical health (e.g. hemi paresis, arthritic hands), occupation or leisure activities (e.g. drumming, piano playing or typing may mean heightened ability to perform rhythmic hand movements), education level and mood state can impact on performance scores (Eastvold et al., 2007).

Despite its apparent simplicity, the BDS scoring system is reliant upon administrator judgement, in particular item 9 which is a subjective measure of insight. It is suggested that the quality of insight can be clinically important in identifying diminished capacity for independent living. Although the measure requires subjective judgement from the administrator, this item has been shown to have very good inter-rater reliability for trained administrators (rho = .88) and has been highly correlated with overall BDS scores (r = .73).

High inter-rater reliability was found when both test administrators were experienced. Less experienced administrators show lower but acceptable levels of reliability (Grigsby & Kaye, 1996). Standardised administration is therefore very important to accuracy of results and ecological validity. Other weaknesses noted by Eastvold and others (2007) were a limited scoring range which can predispose to ceiling effects, particularly in a non-geriatric population on the go/no-go tasks. The subjective nature of the scoring system and limited possible scores on the MBDS were given as the main likely reasons for the current limited application in clinical healthcare settings.

Ability to Predict Outcomes

The BDS has value in subjectively predicting the capacity for self-regulation, motivation, planning, task shifting, monitoring and updating of relevant information in working memory, and execution of intended goal directed behaviour (Lezak, 2004; Suchy & Bolger 1999). The traditional manual version of the BDS has well established validity and in predicting the ability of geriatric patients to function independently. High ecological validity is stated by the authors as differentiating the BDS from other executive measures in that it predicts a person's capacity to regulate their own behaviour in an unsupervised context. The manual version of the BDS has been used in the elderly population to predict functional autonomy and plan appropriate levels of rehabilitative care following hemiparesis and other chronic medical conditions. Hall and Harvey (2008) used the BDS to examine possible differences between samples of patients with Alzheimer's and Vascular Dementia in cognitive functioning. It was reported that performance on problem-solving tasks requiring the voluntary control of behaviour could help to differentiate between mild cognitive impairment and normal ageing. Myers, Grigsby, Teel, and Kramer, (2009) examined the addition of the BDS to nursing assessment of functional autonomy capacity in patients admitted to a rehabilitation unit following fractures, medical and surgical conditions. The study concluded that using the BDS to indicate level of EF enhanced the predictive accuracy of the existing model being use, and suggested,“the inclusion of this measurement of executive cognitive function may significantly enhance the accuracy of nursing prognoses " (p.264).

The utility of the BDS has been shown in various patient populations in a range of contexts. A study by Suchy and others (2003) compared the clinical utility of the BDS to a range of traditional measures of EF in a non-geriatric population. The researchers were interested in differentiation of mild, moderate or severe classification of patients who had sustained a TBI and lesion location. The result showed "the Fluid Intelligence Factor of the BDS improved classifications above and beyond traditional measures" and in particular was sensitive to detecting those who had sustained mild injury, and to frontal lobe integrity (p.492). The Fluid Intelligence Factor was defined by the authors as measuring working memory and insight. Traditional measures with proven sensitivity to global executive dysfunction used for comparison were the Trail Making Test-B, the Stroop Colour Word Test and the Controlled Oral Word Association test. In combination these were considered equivalent to the BDS in administration time and difficulty. The BDS was shown to successfully classify participants in terms of severity and location of injury demonstrating sensitivity to detecting subtle executive chronic deficits following TBI. Traditional measures have been shown to be relatively insensitive to mild injury to the ventral frontal region (Eastvold et al., 2007). Behavioural, social and emotional changes are often precipitated by TBI (Spitz et al., 2012). Therefore means of predicting the likely extent of these changes and associated difficulties could usefully inform early intervention and treatment strategies. Although factors such as age, educational history, pre-injury productivity, and injury severity are closely associated with predicting level and speed of rehabilitation following TBI, an individual's cognitive abilities have been shown to predict functional outcomes (Spitz et al., 2012). Furthermore according to Spitz and others (2012), the domain of EF provided a stronger prediction of this over the first year post injury than demographic and injury severity variables. The authors concluded that executive functions were more strongly related to functional outcome than processing speed or memory.

Belanger and others (2005) examined the BDS as a measure to assess motor and cognitive regulation in patients with mild cognitive impairment and Alzheimer's disease. The author's found that performance on the BDS to be an independent predictor of Activity of Daily Living (ADLs) in both these groups. High ecological validity is stated by the authors as differentiating the BDS from other executive measures in that it predicts a person's capacity to regulate their own behaviour in an unsupervised context as referred to in the in the earlier section on AJ Luria and his work highlighting the importance of ecological validity in neuropsychological assessment. This relates to outcomes across a range of disorders from TBI to chronic health conditions such as diabetes, heart disease, psychological disorders and other serious medical conditions in which long term medication needs to be self-administered. Complex organisational and management processes are required to plan and self-monitor a personal medication regime on a long term basis. According to Stilley and others (2010) approximately 50% of those with chronic disorders experience difficulty in adhering satisfactorily to their schedule.

These findings support Luria's theory and the hypothesis that the wider availability of high quality neuropsychological assessment in routine patient care settings could supplement clinical evaluation enhancing outcomes for patients and families (Achiron, Doniger, Harel, Appleboim-Gavish, Lavie & Simon, 2007). Computerised testing has the potential to assess the subtle aspects of EF into a more quantitative, measurable and comparable data form than is possible within the existing parameters of the manually administered version. Efficacy and criterion validity of the manual version of the BDS have already been demonstrated in chronic medical and TBI rehabilitation settings; however it seems reasonable to assume the availability of an efficacious computerised version would have wider healthcare applications because it would have a greater degree of clinical and practical utility.

Development of the Computerised Behavioural Dyscontrol Scale

Computerised versions of the BDS have been evolving over the past four years, potentially seen as highly beneficial and useful in a wide variety of healthcare settings, however further research is needed to establish validity, accuracy and reliability of these measures. The computerised BDS was developed by Robert Kooken in collaboration with Jim Grigsby and is closely based on the manual version designed by Jim Grigsby and Kathryn Kaye (1992). Items on the computerised version were designed to replicate items on the manual version as far as possible. Precise equivalence was not achieved because of different modes of administration; however, Luria's (1980) original ideas about means of examining different components of behavioural control were used. Given the CBDS was developed by the same people who designed the MBDS, there was consistency in the design of task requirements and overall conceptualisation of the computerised version. Tasks were aimed at measuring a person's capacity to effectively initiate and sustain purposeful goal directed behaviour, and inhibit inappropriate activity (Grigsby & Kaye, 1996). Standardisation, accuracy and consistency of information recording and presentation were expected advantages of this mode of administration.

The test being examined in this study is preceded by previous research into the efficacy of an electronic version of the Behavioural Dyscontrol Scale was performed in 2007 by Eastfold, Suchy, Strassberg,Wilson, and Whittaker, also by Cope, Derbridge and Suchy (2005). This research was the first to investigate the possibility of converting the traditional BDS into an electronically administered measure and aimed to widen the clinical utility of the BDS. Cope et al., (2005) researched the efficacy of the BDS-EV in a sample of fifty five community dwelling adults between the ages of 18-68 years. Results from the electronic version were compared to scores on a battery of traditional clinical tests of EF (Stroop word/colour test, Trail making-B, WAIS-111 Information subtest, Ruff figural fluency test, Controlled word association test). A bespoke electronic response console was designed from which the tasks were performed using large arcade style response buttons and a joystick. The developers aimed to reduce the extent to which the BDS-EV relied on language and memory describing these as non-executive processes. Therefore, both verbal and written instructions were given to participants, and most tasks had practice trials. Seven tasks of varying difficulty levels, generating 15 variables, meant that more information could be elicited whilst floor and ceiling effects could be minimized. As with the manual version, the BDS-EV aimed to measure three components of executive functioning: working memory, capacity for inhibition and response selection/set maintenance (Cope et al., 2005). Good internal reliability and consistency was shown with Cronbach's alpha values ranging from.70 to .87. The BDS-EV total score was found to be "reasonably comparable" to the BDS manual version total score (Cope et al., 2005 p.22). Construct validity correlated highly after age and processing speed were accounted for. Support for the incremental validity of the BDS-EV was also found. Eastvold and others (2007) concluded that the computerised administration improved sensitivity to subtle executive weaknesses. This supported previous findings that "traditional measures of executive functioning are generally insensitive to subtle executive deficits associated with mild head trauma" (p.69).

Overall, similar time was required for administration of the BDS-EV as with the manual version, however administration was easier and results provided for increased accuracy of participant classification (Cope et al., 2005). Seven tasks were developed which together generated 15 variables, significantly increasing the number from 9 on the MBDS. Areas of difference between the two methods of administration included: the BDS-EV contains measures of speed, whereas the BDS manual version is untimed apart from the letter/number sequencing, the differentiation of individual abilities are weighted differently among items, and the BDS manual version contains a separate item for insight measurement, whereas the BDS-EV does not contain a specific item for this. The authors conclude that whilst showing promise, the BDS-EV in its present form is not ready for clinical use. Further research and validation within a patient population using larger sample sizes is recommended. It is suggested that the two instruments “should not be viewed as interchangeable”. It is suggested that the BDS-EV be viewed as a new measure “inspired by, rather than being psychometrically parallel to the manual BDS” (Cope et al., 2005, p. 24).

The tasks comprised two baseline items, one finger tapping and one choice reaction time. A Push-turn-tap-tap (PTT) task was developed to parallel four motor programming items from the MBDS in which participants were required to learn four different sequences of hand movements using a “joystick” on the response console. The Ding-tap (DT) was a direct equivalent to the inhibition of mirroring response tasks on the MBDS in which participants were required to do the opposite number of taps to the examiner. Environmental independence and response inhibition were measured using two Go/No-Go tasks where responding to colours and shapes was required. Finally alphanumeric sequencing (ANS) was used as a direct equivalent to the oral item on the MBDS, together with another purer measure of working memory not dependent on visual scanning or motor speed, which required participants to identify objects appearing earlier on the screen from two trials back (2-Back).

In summary, a computerised version of the BDS has potential for being more sensitive and reliable than the manual version. The computer version of the BDS developed by Robert Kooken is an attempt to further refine the BDS-EV into a more practical test with good ecological validity. This may provide us with improved insight into problems which can then be targeted by rehabilitation and hence lead to better outcomes.

Summary of Introduction and Study Aims

In summary traditional EF measures have been criticised for a lack of 'real life' situational interference. The BDS includes tasks that test the capacity to appropriately select and implement attentional resources amidst multiple competing environmental stimuli. This provides information about a person's capacity to utilise their cognitive, attentional, self-regulatory, and motor resources to plan and engage in behaviour that is meaningful and satisfactory to them in the context of their life. A measure that indicates the extent to which a person has control over their voluntary behaviour (which may not match their intentions to carry out that behaviour) may be the main feature that differentiates the BDS from other traditional EF measures (Hall et al., 2008). The development of a computerised version that could accurately and reliably assess an individual's potential ability to take action that will enhance their well-being following brain injury or illness such as stroke, diabetes, bi-polar disorder, obesity, or substance abuse, would have many clinical applications in the physical and mental health fields.

The specific aims of this study were:

1. Examine to what extent the CBDS and MBDS measure the same construct of EF based on Luria's Three Functional Units theory.
2. Determine any associations between the CBDS and previously standardised validated measures of EF the IVACPT and Trails B.
3. Investigate what influence gender has on performance on the CBDS.

Methodology

Participants

Participants in this study were 40 tertiary students enrolled at the University of Waikato. Of these 12 were male and 28 were female and participant's ages ranged from 18 to 65, with a mean age of 25.6 years. Thirty four identified themselves as New Zealand European or Pakeha, whilst three identified as New Zealand Maori and one reported Kurdish origin see Table 1. Of the 40 original participants results from 38 were used due to incomplete recording of two data sets. The majority reported moderate to high levels of computer use with an average of 37 texts sent per day. Only two people were left- handed, the majority did not smoke and reported consuming occasional to regular amounts of alcohol. All participants were tertiary students and spoke English as their first language.

Table 1 ***Age, Education and Ethnicity of Participants***

	Male (n=12)	Female (n=26)	Overall (n=38)
Age (years)			
Mean	28.08 (12)	24.50 (26)	25.63 (38)
SD	14.196	9.705	11.243
Range	18-65	18-55	18-65
Education (years			
tertiary)	2.500	1.860	2.068
Mean	2.1532	1.186	1.560
SD	1-7	1-6	1-7
Range			
Ethnicity			
% NZ European	83.3 (10)	92.3(24)	89.5(34)
% Maori	8.3 (1)	7.7(2)	7.9 (3)
% Kurdish	8.3 (1)		2.6 (1)

Measures

Demographic Questionnaire

The questionnaire (see Appendix D) was used to obtain demographic information about the participants that could potentially influence performance on measures used. Computer use, amount of texting per day, left or right handedness, sight and hearing, educational level, previous head injury, anxiety, depression and current medications were asked about as they have been shown to alter performance on computerised psychological tests in previous studies.

The HADS (Hamilton Anxiety and Depression Scale)

The Hospital Anxiety and Depression Scale is a 14 item self-report measure designed to screen for anxiety and depression and the relative severity of each as experienced over the past week. This test was used because it assessed both anxiety and depression, was simple and brief to administer. Reliability was demonstrated by Cronbach's alpha of .80 to .84 for Anxiety and .71 to .84 for Depression. It was asserted that the HADS screening is an efficacious measure, although not a diagnostic measure, in non-clinical populations. (Spinhoven 1997; White 1999). Individual scores for anxiety and depression are produced with a maximum potential score of 21 for each (Snaith & Sigmond,1994). A score of 11 or more on either scale is regarded as clinically significant. In this study any participant scoring 11 or more was referred to the student counselling services at Waikato University.

Integrated Visual and Auditory Continuous Performance Test (IVACPT)

The IVACPT designed by Sandford and Turner (1994) is a standardised combined auditory and visual computerised continuous performance test providing objective data about a person's attentional capacity and impulse control. The software developed by Brain Train is designed to detect impairment rather than superior performance. The measure is primarily used to assist in the diagnosis of ADHD, but also in the detection of a variety of disorders related to attentional and self-control difficulties (Strauss et al., 2006). Normative data is from 781 subjects across ten age brackets and both genders. Research shows high reliability and validity in diagnosing and assessing ADHD in children, adolescents and adults (Sandford, 1994).

This measure was selected because it could provide specific accurate measurements about participants' speed of response, variability of response, set shifting ability, sustained attention and response inhibition capacity which could be compared with the computerised BDS output measures. Because the manual BDS is scored using an ordinal score based primarily on subjective observation of performance, the IVACPT provided another measure accurately scoring time sensitive variables in a consistent and standardised way that could be used for comparisons with the computer BDS.

The task was administered using a standard laptop computer with a mouse. Test instructions were both audible and visual. Task requirements involved the pseudorandom presentation of the numbers '1' and '2' alternating between visual and auditory modalities. Impulsivity and inattention are measured both separately and simultaneously over twenty minutes. The subject is required to respond only to the target '1' and inhibit responses to the non- target '2' by clicking on the mouse. Frequency of presentation of targets is varied; typically higher omission errors occur when targets are presented infrequently because demands on attention capacity are greater, and higher commission errors occur when targets are presented frequently because demands on impulse control are increased (Strauss, Sherman, & Spreen, 2006). Score labels are based on a positive interpretation of performance, for example the term 'vigilance' is used rather than 'inattention'. The process produced the following categories and sub categories of scores for each participant:

Full Scale Response Control Quotient (FSRCQ) – composed of:

- Auditory Response Control: Prudence (ARCP), Consistency (ARCC), Stamina (ARCS).
- Visual Response Control: Prudence (VRCP), Consistency (VRCC),Stamina (VRCS)

Full Scale Attention Quotient (FSAQ) – composed of:

- Auditory Attention: Vigilance (AAV), Focus (AAF), Speed (AAS).
- Visual Attention: Vigilance (VAV), Focus (VAF), Speed (VAS).

Sustained Visual Attention Quotient (SVAQ)

Sustained Auditory Attention Quotient (SAAQ)

From the range of scores generated, we selected the following:

Full scale Response Control Quotient (FSRCQ) composed of subscales, Prudence, Consistency and Stamina which measure different aspects of impulsivity. *Prudence* (VRCP, ARCP) subscales measured errors of commission invited by presenting long segments of either stimuli '1'or '2' followed by the alternative stimulus. The *Full scale Attention Quotient (FSAQ)* primarily measures global attention capacity. *Vigilance* a sub category of FSAQ (VAV, AAV) assessed inattention as evidenced by omission errors which were prompted by prolonged segments of non - presentation of either stimulus. *Speed,* FSAQ (AAS, VAS) assessed mean reaction time throughout the whole test reflecting mental effort and information processing speed. We were also interested in the Sustained Auditory and Visual Attention Quotients (SAAQ and SVAQ) because they are global measures of a person's ability to make reliable, fast, accurate responses to stimuli under low demand conditions and then adapt to more demanding conditions demonstrating cognitive flexibility. For our final analysis we used the two global scores as our dependent variables, the FSRCQ and the FSAQ because these totals represented the composite sub-scores itemised above for impulsivity and attention.

Trail-Making Test (TMT)

The Trail-making test was originally developed by Partington and Leiter (1949) to measure complex visual scanning, motor speed, attention, executive function, and in particular cognitive set shifting ability (Lezak, 1995; Stuss et al., 2001).The first part of the test (Trails A) requires the participant to connect consecutive encircled numbers which are presented randomly spread over an A4 page, with a continuous line. The second stage (Trails B) involves the same procedure but alternate numbers and letters have to be connected making the task more demanding. Both parts A and B include a demonstration by the administrator. Completion takes five to ten minutes. Time taken to complete each section was the measure of interest.

Its brevity, ease of administration and scoring together with evidence showing no meaningful differences in scores as a function of ethnicity or gender have made it a commonly used test in neuropsychological assessment. High levels of validity and reliability have been found, however it has been shown to be subject to some practice effects, with performance increasing in adolescence and declining after age thirty (Reynolds, 2002).

This measure was selected because the literature shows it correlates well with other measures of attention ability and processing speed, for example the Symbol Digit Modality Test, and the Wisconsin Card Sorting Test (Reynolds, 2002; Strauss et al., 2006). According to Strauss, Sherman and Spreen (2006) there is some evidence that Trails B has particular use in the assessment of executive functioning and can be predictive of everyday adaptive functioning capacities.

Performance has been found to be affected by age with accuracy rather than speed declining. Lower levels of educational achievement and poorer IQ are associated with poorer performance particularly on part B (Frais et al., 2006). Gender and ethnicity have been shown to have little impact on scores. Test re-test reliability has been shown to be adequate to high with reliability co-efficient of .79 (part A) and .89 (part B) in a sample of neurologically stable adults aged 15-83 years. Practice effects were evident but only for short retest intervals. Inter-rater reliability has been reported as .94 for Part A and .90 for Part B (Strauss, Sherman & Spreen, 2006). Mean completion time for Part A in a sample of neurologically stable adults was reported as 26.52 seconds (*SD* 11.66), and for Part B 72.05 seconds (*SD* 45.22). Parts A and B only correlate moderately well with each other (r =.31-6) suggesting partially different aspects of executive functioning may be measured. Strauss et al., (2006) suggests that Trails B makes greater demands than Trails A on visual search and motor speed because there are more items a greater distance apart, and on cognitive flexibility and ability to maintain set because of the alternating letters and numbers.

The first part of the test (Trails A) requires the participant to connect consecutive encircled numbers which are presented randomly spread over an A4 page, with a continuous line. The second stage (Trails B) involves the same procedure but alternate numbers and letters have to be connected making the task more demanding. Both parts A and B include a demonstration by the administrator. Completion takes five to ten minutes. Time taken to complete each section was the measure of interest.

Behavioural Dyscontrol Scale Manual Version

As described in the introduction, the BDS is a brief, standardised and easy to administer measure, designed specifically to assess the component of neurological functioning that organizes and regulates goal oriented self-directed behaviour. Standardised administration is very important to the accuracy of results and ecological validity. Participants were asked to perform tasks as smoothly, accurately and quickly as possible as they could after watching a short demonstration by the administrator. They were advised that some of the tasks they were about to see may look deceptively easy however it was common for some people to find some parts harder than they expected. Tasks were administered as described in the manual and a brief description what each of the nine tasks involved is outlined in the following paragraphs.

The first task requires participants to tap twice with the right hand and once with the left on the desk surface until they are told to stop. A short practice time is allowed before the actual performance is assessed for smoothness, speed and accuracy.

The second task requires participants to perform the same action changing hands. This task measures the capacity to shift attention from one set to another, working memory and information processing speed, initiation and control of appropriate responses.

The third task requires participants to squeeze the administrator's hand on hearing the word "red" and do nothing on hearing the word "green" (fifteen repetitions). Stimuli were presented randomly including both consecutive and alternate order. This item tests response inhibition capacity. It can be difficult to inhibit the impulse to squeeze on every stimulus which can produce a high frequency of inappropriate responses to the word "green" immediately following a correct response to the word "red". This kind of perseverative error can also occur when the participant squeezes more than once in response to the word "red". Delayed responding to the word "red" can indicate compensation for the tendency to respond impulsively or effort required to initiate purposeful behaviour. The regulation of motor action by the use of inner speech is also tested by this item. The ability to self-correct and errors of dysinhibition and initiation were noted by the administrator.

Task four also requires inhibition of a mirroring response. When the administrator taps the table twice, the participant is required to tap the table once. If the administrator taps the table once, the participant taps the table twice (ten repetitions). As in the previous task, stimuli were presented in a random order including both consecutive and alternating repetitions. Errors were counted for ten repetitions of the task. Echopraxic (repeating the same rather than the opposite movement from the administrator) and perseverative errors were most likely on this task. Delay in responding was also observed as this can indicate compensation for a tendency to respond impulsively or difficulty in information processing. Patterns of difficulty throughout the task were noted as well as participants speed and capacity for self- correction.

Task five requires participants to touch the table alternately with the thumb and each finger as fast and smoothly as possible (five full repetitions after allowing a practice). This task assesses procedural learning capacity. Effortful performances are expected initially with increasing automaticity over several trials.

Task six involves alternately placing a clenched fist, the edge of the hand and the palm of the hand on the table (while saying the words “fist” “edge” “palm” out-loud) and repeating the sequence of actions. This task measures procedural learning and a normal performance would be expected to rapidly improve with practice requiring less deliberate effort and becoming increasingly automatic. Luria observed on this task that using the clenched fist position flat on the table is more difficult than placing the fist down with the thumb on top of the fist. Mis-positioning of the fist was the most commonly occurring error in this task.

Task seven referred to as the HEAD'S Test requires the participant to sit opposite the administrator and copy a series of five hand and arm movements without mirroring. Movements performed by the administrator consist of the following: left fist beside head, right index finger pointed to right eye, left hand vertical, right hand horizontal forming a "T", right hand with bent finger under chin, left hand to left ear. The first mirroring error should be corrected although still scored as an error. Those lacking insight may continue to repeat the same mistake. Self- correction and participants comments on their own errors may provide information for scoring item 9 (insight). This task reflects capacity for inhibition of reflexive responses and echopraxic (mirroring) responses are the most likely type of error.

The Alphanumeric Sequencing (ANS) task requires participants to complete the sequence 1A2B3C4D5E6F7G8H9I10J11K12L verbally, as fast and accurately as possible from memory. This item is timed with a stopwatch and a cut off-score of 20 seconds used for the purposes of scoring. Time taken on this item is the dependant variable of particular interest. This task measures the capacity to shift attention from one set to another, working memory, information processing speed and control.

The final task (nine) is not administered to the participant but represents the administrator's assessment of demonstrated insight into the individual's own performance including frequency and severity of errors. Observations were made throughout all previous items of how the participant reacted to mistakes made and levels of insight and self- awareness expressed. The presence of anxiety in response to errors may indicate an awareness of deficits and the possibility these may be significant. The accuracy with which the participant reports on their performance indicates capacity for insight. Failure to comment on a deficient performance can also be significant. For the purposes of this test the assessment of insight can be divided into different levels.

These component levels for scoring this item:

- Awareness of the *existence* of a deficit
- Awareness of the *nature* of the deficit
- Awareness of the *severity* of the problem
- Awareness of the *significance* of the problem

Although the measure requires subjective judgement from the administrator, this item has been shown to have very good inter-rater reliability for trained administrators (rho = .88) and has been highly correlated with overall BDS scores (r = .73).

The BDS scoring system rating performance on each item from 0-3 was used for the purposes of this research. Performance on all tasks was carefully observed by the researcher and recorded for scoring on the rating scale of 0-3 according to instructions in the BDS manual. Items were subjectively judged and given a score ranging from 0-3; 3 being no errors and a smooth performance requiring little effort, 0 being completely failing to learn the task. Notes were taken from observing the performance for smoothness, number of errors, speed and accuracy, spontaneous correction of errors and participants comments about their own performance. A score for each task was given when all items had been completed. The individual item scores were summed to form the overall BDS score for each participant. The maximum achievable is a score of 27 (BDS Total Score). Among a younger neurologically high functioning group such as our sample, higher scores would be expected. As described earlier, previous studies have used the three categories of Fluid Intelligence Factor (FIF), Motor Programming Factor (MPF), and Environmental Independence Factor (EIF) to subdivide the scores. The MBDS consists of nine items that closely represent Luria’s Three Functional Units (see Table 2 below). Four motor programming tasks (MPF) examining the ability to volitionally initiate and sustain appropriate responses; Two go/no-go tasks (EIF) designed to assess environmental independence and dependence (response inhibition and impulsivity); Three working memory tasks and insight items (assessing fluid intelligence and ability to reason and use feedback) as demonstrated in the following table.

Table 2 ***Manual BDS Tasks Mapped onto Luria's Three Functional Units***

Functional Unit	FIF	MPF	EIF
BDS Tasks	RRL/LLR (1,2) ANS (8) Insight (9)	RRL/LLR (1,2) Alternate Finger/Thumb touching (5) Fist-Edge-Palm (6)	Red/Green "squeeze" hand (3) 1Tap/2Tap (4)

Note:

FIF= Fluid Intelligence Factor, MPF = Motor Programming Factor,

EIF = Environmental Independence Factor, (1-9) = Task order of administration

Computer Behavioural Dyscontrol Scale

A standard Del computer screen and keyboard were used, with stereo speakers and a mouse. The test comprised seven separate tasks which had all been developed ready for administration, however what data was to be extracted had not been programmed so it was necessary to decide this in collaboration with a computer programmer for each task. Instructions were given via an onscreen person (Kathryn Kaye) demonstrating and/or describing each activity and what was expected of the participant. In some tasks, visual cues were used to indicate start and stop instructions. A variety of visual and auditory cues were used at different items, these were delivered by the computer as on-screen messages or auditory signals. Participants were instructed to work as fast and accurately as they could on each task and each task has a practice session. Responses were all performed using the computer keyboard keys. Completion of these items took 15 minutes.

The following seven items from The Digital Executive Battery software (Robert Kooken, Jim Grigsby & Kathryn Kaye) were then presented to participants:

Task 1- FJ finger tapping alternate LT and RT index fingers

Participants press the F and J keys alternately with the index finger until the words “stop” and “press the spacebar to continue” are presented visually on the monitor after 30 seconds. The following data was extracted from this task: total number of keystrokes, total number of correct sequences, number of correct sequences prior to first error, total number of incorrect sequences, mean response time for correct sequence, total number of 1, 2, 3, and 4 category errors (errors categorised as follows: extra character F or J, incomplete sequence, character error other than F or J). The specific dependent variables used for analysis to create a Total CBDS Score are shown in Table 3 at the end of this section).

Task 2 – finger tapping JJF and FFJ

This computer item requires the participant to tap the J key twice immediately followed by the F key once with the index finger of the dominant hand as fast and rhythmically as they can until they are told to stop. After 30 seconds the words *“stop”* and *“press the spacebar to continue”* appears on the screen to continue. The following dependent variables were recorded : total number of keystrokes, total number of correct completed sequences, number of completed sequences before the first error, number of incorrect sequences, number of each category of error1, 2, 3, 4. Errors types were categorised as follows: 0 = no error, 1= extra character of F or J perseveration (for example F , F, F, J), 2 = incomplete sequence (for example F, J), 3= character error other than F or J (for example F, Y, J), 4= end of file (for example F, F, J, F, J time ends). Other variables measured were, response time intervals between keystrokes, fastest sequence, length of time of each correct sequence in order from first response, mean correct sequence response time, and mean incorrect response sequence time. Dependent variables selected for analysis are shown in the table at the end of this section in Table 3.

Task 3 – Spacebar Red / Green Inhibition (Go/No-Go)

Participants are asked to listen to the computer commands. If they hear the word “red” they are to press the spacebar once as fast as they can. If they hear the word “green” they are to do nothing. This continues for 40 commands. Dependent variables measured were: number of correct responses, number of incorrect responses, position in sequence of first error, number of correct red command responses, number of correct green command responses, number of red command errors, number of green command responses , perseverative errors A (incorrect spacebar press on the “green” command following immediately after a correct spacebar press on a “ red” command), perseverative errors B (an incorrect non- response to a “red” command following a correct non- response on a “green” command), individual response times for spacebar presses, mean response times for correct red responses (spacebar press), and mean response times for green incorrect responses (spacebar press).

This item measures capacity to supress reflexes, adapt motor behaviour to changing environmental stimuli, inhibit inappropriate responses, impulse control, attention and self-regulation. Number of correct responses and mean response time for spacebar presses were used as dependent variables as shown in the table 3.

Task 4 – Spacebar response to 1 tap or 2 taps

This is a go-no/go task. The participant is asked to press the space-bar once when they hear the computer make two taps, if they hear the computer make one tap they have to respond by pressing the spacebar twice. These instructions are then presented on the monitor in written form. The participant presses any key to begin. The task lasts for 1 minute 50 seconds. This task is similar to the previous task measuring response inhibition, impulsivity, and errors of commission (perseverative errors).

We measured the following dependant variables: total number of correct and incorrect responses, number of responses before the first error,1 tap responses correct (2 tap command), 2 tap responses correct (1 tap command),1 tap errors (2 tap command), 2 tap errors (1 tap command), perseverative errors, mean response time 2 tap command, mean response time 1 tap command.

Task 5 A-L sequence

This is a warm up exercise for the next task. The participant is asked to type the letters of the alphabet in order from A to L working as quickly as they can. They are instructed to press the spacebar to begin and when they are finished. For this task we measured the following dependent variables:

number of correct keystrokes, number of incorrect keystrokes, place of first error, total sequence time, time from A-F, and time from G-L.

Task 6 – Letter number sequencing

Participants were instructed to type the sequence 1A2B3C4D5E6F7G8H9I as fast and accurately as they could for 30 seconds until they could see the “stop” sign appear on the screen, using only their index finger. The task is a measure of working memory (set shifting, recall and recognition), attention, cognitive flexibility inappropriate response inhibition (impulse control, self - regulation) and appropriate response selection (information processing, modification of response patterns). It also provides good indications of self-awareness and insight. The participant’s comments about their own performance can be examined in relation to their actual performance as evidenced by response times and number of correct sequences. Given it was important to reflect the participant’s performance as accurately as possible in real terms a manual visual search of the data was used to identify errors. The error coding and final list of dependant variables to be measured were as follows: total of errors, perseverative errors (repetition of any digit or character inappropriately), sequencing errors (out of sequence digit or character), non-sequence digits or characters, completed sequences (those that get to 9 or I), correct sequences, incorrect sequences, incomplete sequences, sequence number of first correct sequence, mean digit response time, mean character response time, and mean time for correct and incorrect sequences. Time of the first correct sequence was the closest direct equivalent to the manual ANS task and Trails B, both of which measured performance on a discrete sequence. The other selected dependent variables are itemised in Table 3 at the end of this section.

Task 7 – BWKPDM

Participants are instructed to type the sequence BWKPDM. The letters are presented visually centred on the screen for 1 minute. Participants are instructed to keep typing the sequence until the numbers no longer show up on the screen. There is no spoken instruction to stop this task or indication to the participant for how long they need to continue repeating the sequence. The dependant variables measured were: number of correct sequences, number of incorrect sequences, and mean number of errors. Proceduralisation, inhibition of inappropriate responses and selection and implementation of appropriate motor responses are measured by this item. It is a motor programming task and performance would be expected to increase in speed and accuracy with repetition.

Sub-scores from each item were selected for final data analysis and a Total CBDS score created for comparison to the Total MBDS score.

The dependent variables shown in Table 3 below were used to formulate a composite standardised Total score to represent each participants overall performance on the CBDS. These scores were selected based on EF domains developed by Luria and used by the developers of the Manual BDS. When we had the same dependent variables as those used in analysis of the BDS –EV (Cope et al., 2005) these were also included in producing a composite score that would accurately represent performance. Table 3 below shows selected specific scores that we considered best represent each domain based on the rationale given above. Due to the overlapping and interdependent nature of EF processes it was hard to definitively separate the Three Functional Units and separate specific scores measuring each, consequently where necessary some of the scores have been used across more than one domain. This followed Luria's principle of pluripotentiality and was aimed at best representing each person's performance as accurately as possible.

Table 3 ***Specific Scores Selected for the Total CBDS Score and EF Processes Measured by each***

Task	Scores selected	EF Processes measured
RL/LR RRL/LLR	Total correct sequences	Sustained attention Proceduralisation
	Mean response time	Ability to execute goal directed behaviour in the absence of environmental cues
	Fastest sequence	Effort in processing, represents best most automated performance achieved during the block
	Total errors Number incorrect sequences	Self-correction, errors made while learning
Red-Green Go/No - go	Total correct responses Total errors Mean response time	Dysinhibition / Impulsivity Inability to inhibit responding to a previously reinforced stimulus and inhibit mirroring response
1 Tap/2 Taps Go/No go	As above	As above
ABC	Total sequence time	Execution of goal directed behaviour
Alphanumeric Sequencing	Time of 1st correct sequence Number of correct sequences Total errors Mean digit/character response time	Cognitive flexibility Set shifting Processing effort/speed Self- correction Attention Track details of past performance Insight
BWKPDM	Number of correct sequences	Proceduralisation Sustained attention

These individual scores were then mapped onto Luria’s Three Functional Units as shown in the Equivalence of Measures section (page Table 4) for further comparison.

Software Development for the Computer BDS data extraction

All participant responses (i.e. keystrokes) were recorded using software written to run in Super Lab 4. As previously mentioned, the tasks had all been written however which data to extract from the vast amount produced needed to be decided. The data comprised participant code number, task command, response times in milliseconds, and which keys had been pressed. The quantity and complexity of the data produced meant that formulating it in a manageable, useful and meaningful way was a challenging task. This involved two levels of process; the first was deciding specifically what information was needed from each task and then explaining this to the programmer who wrote codes to extract the data. The technical aspect of designing a custom made software programme capable of achieving these specified output goals was complex. We looked at previous research on computerised testing of EF and which dependent variables had been used, then focused on response times, errors, and number of correct responses and liaised with the software technician. This process took around ten one hour meetings. The result of these meetings was a balance between technical possibility and ideal desired aspects of the raw data. A refined, realistic, time appropriate, re-producible, and ‘human friendly’ final version of each participant’s data was produced.

In summary the Digital Executive Battery software (Robert Kooken, 2010) was used for the CBDS to generate a list of sequential event data which was interpreted by a custom data parser. The researcher worked closely with a programmer who used C# and Visual Studio 2010 to design a programme to extract summary data which could be used for visual examination and statistical analysis. The programmer created a unique coding system that recorded certain patterns of data that could be quantified as errors or correct responses. Other aspects of performance were recorded such as response times and cumulative number of responses. Given it was important to reflect the participant's performance as accurately as possible in real terms a manual visual search of the data was used to identify errors. The error coding and final list of dependant variables to be measured were: total errors, perseverative errors (repetition of any digit or character inappropriately), sequencing errors (out of sequence digit or character), non-sequence digits or characters, completed sequences (those that get to 9 or I), correct sequences, incorrect sequences, incomplete sequences, sequence number of first correct sequence, mean digit response time, mean character response time, mean time for correct and incorrect sequences (see Table 4).

Procedure

Ethical approval was obtained from the School of Psychology Ethics Committee University of Waikato. Potential participants were recruited by displaying posters on noticeboards in the Faculty of Social Science building University of Waikato. A notice asking for participants was placed on the University's Psychology Café website, and students taking Psychology 103 were offered a 1% course credit for participating in the research following consultation with the course convenor. Other participants were offered a $20 petrol voucher. Age Concern in Hamilton were contacted and a poster (Appendix B) asking for participants and explaining the research was placed on a noticeboard in view of members, however no participants were recruited from this source.

Individuals who were interested in the research contacted the researcher by phone or email and were provided with an information sheet (Appendix C) outlining the project, requirements and time commitment. All participants needed to have completed at least NCEA level 1, speak English as their first language, have no significant hearing or sight impairment, and not be part of any other research on TBI. Of the 50 people who made contact with the researcher, 10 did not meet the eligibility criteria or were unable to attend the session times. Following recruitment, meetings were scheduled in a small room in the Faculty of Social Sciences University of Waikato. Each session began by the researcher introducing herself, reviewing the aims of the research, outlining confidentiality and anonymity, the voluntary nature of the commitment, the participant's right to withdraw at any stage, and the likely session duration of one and a half hours. The nature of the tasks was outlined and participants offered the opportunity to ask any questions. Written consent was obtained. All sessions were conducted with the same desktop computer, using Super lab 4 Digital Executive Battery, (Robert Kooken, 2010) in a small room with no windows. Distractions were minimised by ensuring a quiet environment as far as possible, and cell phones were switched off. The tasks were administered in the following order to avoid consecutive computer tasks and maintain subject motivation as far as possible. The order was as follows:

IVACPT, Trails, BDS (manual), HADS, Demographics, BDS Computerised version. The administration order of the computerised version of the BDS and the manual version of the BDS was alternated in each consecutive trial to counteract order and practice effects as far as possible.

Instructions were given to participants about the tasks in the following statement:

"I'll be asking you to do several different things today, two will be tasks on the computer, three are paper and pencil tasks, one involves asking how you have been feeling over the last week, one is quite fun, the other is a questionnaire. I shall also be asking you to do a number of hand movements after watching me do them, these can look easy but sometimes people can find some of them quite hard, so don't worry just do the best you can. Feel free to ask me anything at any time. Have you any questions now? Are you happy to get started?"

The performance of each participant was observed and scored in accordance with standardised procedures. Participants were thanked for their time and effort and those scoring 11 or higher on the HADS were advised of the student services counselling at the university and how to contact them. Vouchers were signed for by the recipient and those opting for course credits were given the appropriate paperwork. Those participants wishing to be sent the final results of the research were invited to leave their email address with the researcher. The session length for each participant was approximately one and a half hours.

Equivalence of Measures

Luria's cognitive theory of EF pioneered the concept of interdependence between areas of the frontal lobes that can also function independently, creating a complex system of interrelated processes able to operate separately or in combination to regulate, activate and inhibit our behaviour. This understanding of EF has increasingly driven the development of modern neuropsychology testing as already described in the introduction. Luria used the three separate yet interrelated functional units of working memory (FIF), appropriate response selection and initiation (MPF), and inappropriate response inhibition (EIF) as a lens through which to view the construct of EF. The researcher mapped the measures used in this study onto these 3 units for the purposes of analysis. The following Equivalence of Measures Table (Table 4) demonstrates how Luria's theory of Three Functional Units relates to the construction of the BDS items (Manual and Computerised Versions), the Trails B, and the IVACPT and the underlying executive processes associated with each. According to Luria's theory of pluripotentiality each of the tasks outlined in Table 4 overleaf tap a primary functional sub-unit, however will also overlap to some extent with other functional sub-units.

Table 4 ***Equivalent Tasks on the CBDS and MBDS***

BDS 3 index scores	**Fluid Intelligence Factor (FIF)**	**Motor Programming Factor (MPF)**	**Environmental Independence Factor (EIF)**
Luria Functional Unit	Working memory, insight.	Appropriate response selection	Inappropriate response inhibition
CBDS	R/L and RRL/LLR - number correct responses-mean response time ANS - number correct sequences -total errors - mean digit/character response times - time 1st correct sequence-mean overall response time	R/L and LLR/RRL - number correct responses - fastest sequence time - mean response time ANS- number correct sequences- fastest sequence time ABC - time of sequence BWKPDM - number of errors-number of correct sequences	RED/GREEN - number of correct responses -number of errors - number of perseverative errors-mean response time. ITAP/2TAP- number of correct responses, number of errors, mean response time - time of first error ANS- perseverative errors RRL/LLR - total number of errors, sequencing errors and perseverative errors
MBDS	Alternate thumb finger touching, HEADS Alphanumeric sequencing, Insight	RRL/LLR Alternate thumb finger touching, Fist-Edge-Palm. HEADS	Red/Green Go/Non-go, 1Tap/2 Tap, Fist-Edge-Palm HEADS
IVACPT	Full Scale Attention Quotient	Full Scale Attention Quotient	Full Scale Response Quotient
TRAILS A B.	TRAILS A, B.		

Data Analysis

Raw data from the CBDS was visually examined, errors and correct responses identified and coded, then extracted manually and entered into excel. Data from the IVACPT measure was extracted and then entered into excel. All other data was directly entered into excel from where analysis in SPSS was undertaken using IBM Version 20 (2011). Two participant's data sets were incomplete and subsequently removed from the analysis leaving thirty eight individual data sets for analysis. SPSS was then used to calculate participants' scores on each of the measures to create summary data. Outliers were then identified using boxplots, removed from any subsequent analysis and replaced with the mean. The analyses performed from this point are described in the following section.

Results

The main aim of this study was to explore the extent to which the CBDS measures the same construct as the MBDS based on Luria's Three Functional Units. Secondly to get a more complete picture of how the scores generated by the CBDS relate to two other well validated measures (Trails and IVACPT) of EF, and thirdly to asses influence of gender difference on performance.

The total distribution of scores on each measure was tested for normality. Given the data produced from the MBDS was ordinal (with the exception of the timed alphanumeric sequencing task) and the data produced from the CBDS and IVACPT were continuous, parametric and non- parametric statistics were used for the analysis. In order to obtain a total score for the CBDS for comparison to the total score on the MBDS a number of computer generated scores were selected from each task that best represented performance on that item (see Table 3). Scores from the computerised test were then standardised by transforming into z-scores to enable the creation of a Total Composite Computer score for each participant. A high score on the MBDS represented better performance, whereas on the CBDS this only applied to a minority of scores. The specific z scores selected were based on the conceptual understanding of Luria's Three Functional Units of Fluid Intelligence (working memory and insight), Motor Programming (initiating appropriate goal directed behaviour) and Environmental Independence (inhibition of inappropriate behaviour). These three units theoretically represent three different domains of the EF construct as were applied to the development of the original manual BDS. The selected scores were then mapped onto the three indexes of FIF, MPF and EIF (as previously described in the method section) to facilitate comparison with other measures. The scores contributing to each of the indexes are described in Table 3 and are listed in the method section. To obtain accurate results, the z- scores were reversed by multiplying them by minus 1 in cases where a low score represented high performance, before calculating the Total CBDS. This meant that all numerically high scores signified better performance than numerically low scores and all scores representing a good performance went in the same direction.

To directly address the research questions firstly the total distributions of scores on each measure were tested for normality and internal consistency. Inter- correlations between Total CBDS and MBDS scores and Three Functional Unit scores were performed to assess convergent validity and the extent to which the two versions were measuring the same construct. Comparisons were then made between scores on the CBDS and other previously validated measures of EF; the Trails and IVACPT using Spearman's rho and Pearson's r correlation co-efficient. Similar results were yielded,and Pearson's r were reported. To gain a more detailed picture of how much the manual and computer versions were measuring the same construct, scores on individual equivalent items (see Table 4 in measures section) were then correlated. Finally results were reported for the influence of gender on performance across all measures.

Descriptive Statistics

Descriptive statistics for the MBDS and CBDS scores can be found in in Tables 5 and 6 respectively. Descriptive statistics for all individual items used to form the Total CBDS composite score and individual MBDS items are shown in tables at the end of this section in which gender influence is reported.

Table *5 Means, Standard deviation (SD), and Ranges for the Manual BDS Total and Three Composites*

MANUAL	N	Range	Minimum	Maximum	Mean	Std. Deviation
TOTAL	38	15	11	26	21.8	3.43
EIF	38	10.58	7.53	3.05	000	2.45
MPF	38	11.95	-7.51	4.54	000	3.19
FIF	38	13.62	-9.59	4.03	000	3.05

Note:

TOTAL = Total Score Manual Behavioural Dyscontrol Scale, EIF = Environmental Independence Factor, MPF = Motor Programming Factor, FIF = Fluid Intelligence Factor

On the MBDS our population performed at higher levels than a previously tested sample of 43 young adults whose mean Total Score on the BDS was 17.6. The range and standard deviation were similar across FIF, MPF, and EIF domains.

Table 6 ***Means, Standard Deviations (SD) and Ranges for the Computer BDS and Three Composites***

COMPUTER	N	Range	Minimum	Maximum	Mean	Std. Deviation
TOTAL	38	73.55	-42.12	31.44	24.21	14.47
EIF	38	36.12	-26.93	9.21	0000	6.09
MPF	38	25.06	-14.07	10.99	-0023	6.00
FIF	38	32.63	-13.92	18.71	0000	6.46

Note:

TOTAL = Total Score Computer Behavioural Dyscontrol Scale, EIF = Environmental Independence Factor, MPF = Motor Programming Factor, FIF = Fluid Intelligence Factor

The computer generated scores yielded a greater range of scores in contrast to the manual version. This would be expected given the fine sensitivity and range of data in comparison to the subjectively scored manual administration limited by ceiling effects and a scoring range of 1-3.

Table 7 ***Descriptive Statistics for HADS Anxiety and Depression, Trails A/B, and IVACPT***

	N	Range	Minimum	Maximum	Mean	Std. Deviation
HADS-A	38	17	0	17	8.13	3.857
HADS-D	38	12	0	12	4.89	2.939
TRAILS A	38	23.22	11.78	35.00	20.135	5.525
TRAILS B	37	170.38	18.62	189.00	50.805	35.502
FSRCQ	38	65	59	124	93.18	17.239
FSAQ	38	104	23	127	99.6	21.610

Note:

HADS-A = Hamilton Anxiety and Depression Scale (Anxiety), HADS-D = Hamilton Anxiety and Depression Scale (Depression), FSRCQ = Full Scale Response Control Quotient, FSAQ = Full Scale Attention Quotient

Our sample showed mean scores below the cut off of 11 for both clinically significant levels of anxiety and depression. Although levels of anxiety were higher than levels for depression the majority of participants did not have levels of anxiety or depression likely to interfere with cognitive performance. Overall scores on the Trails A and B reflected the high functioning level of our sample. Compared to previously reported norms of 26.52 (SD 11.66) on Trails A and 72.05 (SD 45.22) the population tested in this study performed both trials much faster and with less deviation from the mean. This is not surprising given the nature of our sample. The mean scores on the global IVACPT measures were between 90 and 100, with the mean Attention Quotient being higher than the mean Response Control Quotient. Compared to standardised norms our population performed within one standard deviation (15) of the mean (100).

Internal Consistency Reliability

Higher internal consistency is associated with greater consistency in scores and reflects greater reliability. Internal consistency was examined for the MBDS and CBDS Totals and composite scores using Cronbach's alpha.

Analysis yielded reliability coefficients for the Total MBDS and Total CBDS of .788 and .859 respectively. The alphas across Luria's Three Sub-units were .735 for the CBDS and .856 for the MBDS.

Convergent Validity

In order to examine the degree to which MBDS and CBDS demonstrated a pattern of convergent validity, correlations between Total scores and Three Functional Units were carried out.

Table 8 ***Correlations between Total Computerised BDS Score and Total Manual BDS Score across Luria's Three Functional Units (Pearson's r)***

Measure	TOTAL SCORE	FIF	MPF	EIF
	.645**	.499**	.530**	.384*

Note: * p < .05 ** p < .01
N = 38 FIF = Fluid Intelligence Factor MPF = Motor Programming Factor EIF = Environmental Independence Factor

Table 8 demonstrates moderate to high correlations between Total Scores and all Three Functional Units on both modes of administration. The CBDS and MBDS Total Scores yielded a relatively high correlation in a positive direction at the .01 alpha level as expected. Across the two modes of administration the MPF and FIF sub-units showed stronger relationship with the equivalent subunit than the EIF did.

Table 9 ***Correlations between Three Functional Units on the Computer and Manual Behavioural Dyscontrol Scale (Pearson's r)***

Measure	Computer FIF	Computer MPF	Computer EIF
Manual FIF	.499**	.327*	.587**
Manual MPF	.607**	.530**	.378*
Manual EIF	.457**	.345*	.384*

Note: * p < .05 ** p < .01
N = 38 FIF = Fluid Intelligence Factor MPF = Motor Programming Factor EIF = Environmental Independence Factor

Intercorrellations between Luria's Three Functional Units across computer and manual versions are given in Table 9. As the correlation matrix shows all combinations yielded moderately to high significant relationships in a positive direction. Of particular note was the FIF which demonstrated correlations at the 0.01 alpha level across both other domains.

Correspondence with traditional executive measures

Pearson's r correlations were conducted to examine the degree to which the CBDS assessed similar constructs to other previously validated measures of EF.

Table 10 ***Correlations between the Total Composite Computerised Score, Three Functional Unit Sub-Scores and other Measures of EF***

Measure CBDS	FIF	MPF	EIF	TOTAL
TRAILS A	.622**	.349*	.498**	.544**
TRAILS B	.490**	.363*	.344*	.472**
FSRCQ	.124	.126	.435**	.209
FSAQ	.381*	.202	.505**	.521**

Note:

N = 38. * p < .05 ** p < .01
FIF = Fluid Intelligence Factor MPF = Motor Programming Factor EIF = Environmental Independence Factor, FSRCQ = Full Scale Response Control Quotient, FSAQ = Full Scale Attention Quotient

Table 10 shows moderate correlations were found between the Total CBDS score and Trails A and B, with higher correlations resulting between Trails A and B and the FlF domain (working memory) in a positive direction at the .01 alpha level. The FIF also showed a low to moderate degree of relationship with the FSAQ .The EIF domain from the CBDS moderately correlated with Trails A, B as well as the global attention (FSAQ) and impulsivity (FSRCQ) scores on the IVACPT. The EIF was the only functional unit to correlate with all the other traditional measures of EF used in the study. Timed alphanumeric sequencing items (ANS) across the Trails B, MBDS and CBDS measures yielded significant positive relationships at the

.01 alpha level between CBDS and the Trails B, and between the MBDS and the Trails B.

The nine individual MBDS items were then compared to the CBDS equivalent items (see Table 4).

Results for correlations between equivalent individual tasks on the MBDS and equivalent CBDS are outlined in the following tables and paragraphs in administration order of the MBDS item administration 1-9.

Table 11 ***Items 1 and 2 - Manual and Computer RRL/LLR***

Measure	Total number of errors	Total correct sequences	Fastest sequence time	Mean correct sequence time
RRL MANUAL	.341*	.558**	.442*	.408*
LLR MANUAL	.203	.488**	.572**	.171

***Note*: * p < .05 ** p < .01**

Table 12 ***Item 3 Red/Green Go/ No-Go Tasks on Manual and Computerised Administrations (Pearson's r)***

Measure	Number of Correct Responses Computer	Number of Incorrect Responses Computer	Place of 1st Error Computer	Perseverative Errors on Red Command Computer	Perseverative Errors on Green Command Computer
Go/No- Go Red/Green Manual	.036	.004	-.110	-.260	-.084

***Note*: * p < .05 ** p < .01**

Table 13 ***Item 4 - Manual and Computerised 1 Tap / 2 Tap Tasks (Pearson's r)***

Measure	Total Correct Responses	Total 1 Tap Responses to 2 Tap Command	Total 2 Tap Responses to 2 Tap Command	Mean R/T to 1 Tap Command	Mean R/T to 2 Tap Command
1Tap/ 2Tap Manual	.024	.316	-.068	.033	.093

***Note:* * p < .05 ** p < .01**

Table 14 ***Item 5 - Manual Alternate Finger Thumb Sequencing, Computer BWKPDM, 1 Tap/2 Tap, and Alphanumeric Sequencing (Pearson's r)***

Measures	BWKPDM Total Correct Sequences	Mean 2 Tap Response Time to 1 Tap Command	Number of Completed Sequences Alphanumeric Sequencing	Total Correct Responses to 1 Tap/2 Tap Command
Alternate Finger Thumb Sequencing Manual BDS	.083	-.406*	.375*	.233

***Note*: * p < .05 ** p < .01**

Table 15 ***Item 6 - Manual Fist-Edge-Palm, Computer RRL/LLR and BWKPDM***

Measure	RRL-Total Correct Sequences	LLR-Total Correct Sequences	RRL- Fastest Sequence	LLR- Fastest Sequence	BWKPDM
Fist - Edge- Palm	.534**	.392*	.450*	.514**	-.061

Note: *** p < .05 ** p < .01**

Table 16 Item 7 ***- Manual HEADS and Computer Go/No-Go and RRL/LLR Errors on CBDS (Pearson's r)***

Measure	Total 1Tap/2 Tap First Error	Total 1Tap/2 Correct Responses	Red Green Total Correct	Number of Errors RRL	Number of Errors LLR
HEADS Manual	.077	-.127	-.028	.456**	.363*

Note: *** p < .05 ** p < .01**

Table 17 ***Item 8 - Manual Alphanumeric Sequencing and CBDS (Pearson's r)***

Measure	Mean Correct Sequence Time	Time of 1st Correct Sequence	Mean Digit Response Time	Mean Character Response Time
1a2b3c4d5e6f7g8h9i Manual				
Score	.262	.217	.312	-.080
Time	.295	.315	.359*	-.247

Note**: * p < .05 ** p < .01**

Table 18 ***Item 9 - Manual BDS Insight and Alphanumeric Sequencing CBDS***

Measure	Number of Correct A/N Sequences	Time of 1st Correct Sequence	Mean Character Response Time A/N Sequencing
Insight score MBDS	.542**	.463**	.357*

Note**: * p < .05 ** p < .01**

As can be seen in Table 11 all the scores on the RRL/LLR CBDS item showed statistically significant levels of correlation with the equivalent item score on the MBDS. The two measures of Reaction Time were moderately negatively correlated, whilst Number of Responses and the Total Number of

Correct Responses demonstrated the strongest relationship on this task. The Fastest Sequence time showed the highest correlation to the MBDS on this item.

The Go/No-Go (Table 12) item did not show any significant relationship between the manual and computerised versions. No significant relationship was found between the 1Tap/2Tap manual and computerised tasks.

Table 12 shows alternate Thumb and Finger sequencing on the MBDS positively correlated with the Number of Completed Sequences on Alphanumeric Numeric Sequencing and Total Correct Responses to 1 Tap command. A negative correlation was demonstrated with Mean Correct Response Time to 2 Tap Command on the CBDS. This would be expected given a high score on the Mean Correct Response Time was represented by a low score.

The Fist-Edge-Palm item on the MBDS was found to be significantly correlated with the Total Number of Responses on the RRL and LLR items, the Total Correct Responses on RRL and LLR and the Fastest Sequence and on LLR (Table 15) No correlation was found with the BWKPDM item as might have been expected.

The HEADS (Table 16) item on the MBDS did show a low to moderate significant correlation with the number of incorrect sequences on the CBDS RRL task as seen in Table 7 however no relationship was found between the other equivalent items requiring inhibition of mirroring and perseverative behaviours.

Table 17 shows statistically significant relationships were found between each of the following scores on the Alphanumeric Sequencing CBDS and the Time variable on the manual administration; First Correct Sequence Time, Mean Character Response Time, Mean Digit Response Time, Mean Correct Sequence Time. However, no relationship is demonstrated between the subjective 1-3 score ratings for the manual task and the computerised scores. The CBDS was found to be positively related to the MBDS at the 0.5 alpha level on the Mean Digit Response Time score on this item.

Item 9 (Table 18) demonstrated significant relationships between the subjectively scored Insight item on the MBDS and CBDS sub scores representative of working memory (First Correct Sequence Time, Mean Character Response Time, and Number of Correct Sequences). Insight was counted as a working memory item (as defined by Luria).

Demographic variables

Demographic variables previously found to be linked to performance on EF tests were correlated with Total CBDS scores, specifically anxiety, modelling depression, experience with computers, education level and age. It was not surprising that no significant effect was found between computer experience and performance given 40% of participants rated their computer experience as ‘high’, 45% as ‘medium’ and 15% as ‘low’ no significant effect was found, and all were tertiary students. No significant relationships were found between anxiety, depression, or previous computer use and the Total CBDS score yielded correlation coefficients of.102, .104, and .132 respectively. Anxiety and depression correlated positively with each other as would be expected (.433**). Age however yielded a significant correlation with Total scores on the CBDS in a negative direction of -.423**. Levels of daily texting were also examined using Pearson’s r and showed no relationship to performance.

Given some of the data was non-parametric, both parametric (ANOVA) and non-parametric (Mann Whitney) tests were performed to look at the influence of gender on performance across measures. Both tests produced similar results and the parametric ones are reported for each measure. Mean scores, standard deviations and one way ANOVA using an alpha level of 0.5 were calculated and then analyses performed.

Table 19 ***Descriptive Statistics for HADS, Trails A and B***

Measure	Female Mean (*SD*) N = 26	Male Mean (*SD*) = 12	Total Mean (*SD*) N= 38	ANOVA	Effect size
HADS- Anxiety	9.50 (3.49)	5.17 (3.09)	8.13 (3.85)	F (1,36) = 14.0 p =.001**	.280
HADS – Depression	5.54 (2.94)	3.50 (2.50)	4.89 (2.93)	F (1,36) = 4.30 p = 0.045*	.107
TRAILS B	49.98 (35.36)	52.75 (37.49)	50.80 (35.50)	F (1,35) = 0.46 p = .831	.001

HADS and Trails B mean standard deviation and effect size data for males, females and the total sample is shown in Table19. The Trails B data showed females completed the task on average over two seconds (2.77) faster than males, with a total mean of 50.80 seconds across genders. The difference between speed of performance between genders was not however found to be statistically significant. Table 19 also shows female scores on the anxiety subscale of the HADS were significantly higher than males, and moderately significant effects for reported levels of depression were found. Overall levels of anxiety were much higher than levels of depression.

Tables 20-28 summarise Descriptive Statistics and Effect size for gender across all other measures. Table 20 (Appendix A) shows the MBDS mean, standard deviation and effect size for females (28), males (12) and total sample (38). The data shows overall scores were similar for males and females. The mean total score for the MBDS across all participants was 21.58

out of a maximum 26. Different performance between genders was demonstrated on the timed alphanumeric sequencing task (ANS), with females completing the item on average 2.89 seconds faster than males. This is .12 seconds less than the difference in mean completion time between males and females on the equivalent Trails B item (see Table 4 Equivalent Measures) with female's mean time being faster in both measures. Although females performed this task faster than males gender was not found to be statistically significant in determining results on this task.

Table 21 (Appendix A) shows influences of gender on auditory and visual measures of vigilance (attention) on the IVACPT, with females showing significantly higher levels of both than males.

Table 22 (Appendix A) shows results for the R/L CBDS item. Males made a higher number of total key presses than females, although number correct key-presses was the same as for females. This reflects greater accuracy and consistency of appropriate responding with less impulsivity in female's performance than males performance on this task. Males made more key presses than females and recorded the fastest response time, however males also recorded a statistically significant higher level of incorrect responses.

Table 23 shows that on the RRL/LLR item (Appendix A) females exhibited greater accuracy reflected in a higher number of correct sequences being completed before the first error was made, and as with the previous item males made a statistically significant higher number of incorrect responses than females. Males recorded the fastest sequence time.

Table 24 shows performance by males and females was very similar on the Red/Green Go/ No go task with no significant differences emerging. There was little variation in performance between genders or between participants, and very few errors committed on this task.

Table 25 (Appendix A) shows performance on the 1Tap/2Tap task. Overall number of errors and correct responses were similar between males and females. Response times yielded the most difference, showing men’s mean reaction time to be faster than women’s when the task required initiating a 2 key-press response, and women’s reaction time to be faster than men’s when the task demand was one key-press (this difference was not statistically significant). Both genders got faster over time when a single key-press response was required and slower over time when two key-presses were the correct response.

Table 26 (Appendix A) shows females completed the ABCDEF discrete sequence task significantly faster than males whilst also making fewer errors. Overall females performed with greater levels of accuracy and more speed.

Table 27 (Appendix A) shows females performed with greater accuracy and faster response times on the alphanumeric sequencing task (ANS). A statistically significant difference was found on fastest sequence time with females producing a faster sequence time to males. Overall it is interesting to note that mean digit response times were slower than mean character response times for both males and females.

In summary, overall both the CBDS and MBDS demonstrated very good reliability across the Three Functional Units, with Cronbach’s alpha values ranging from .735 to.859. When comparing scores between the two measures significant relationships were found between the Three Functional Units showing a moderate to high degree of specificity to the factor constructs, as well as between Total scores. Convergent validity was confirmed when compared to existing traditional clinical measures of those constructs. The EIF correlated across all other measures, with Trails A and B yielding the highest correlations across all Three Functional Units, in particular with the ANS manual and CBDS equivalent tasks. The nine individual MBDS items were then compared to the CBDS equivalent items (see Table 4 of Equivalent Items).The majority of individual item scores were found to be significantly related (the RRL/LLR, Finger/Thumb Sequencing, Fist-Edge-Palm, HEADS, Alphanumeric Sequencing, and Insight items). Noticeably the Go/No-go items did not yield any significant correlations. Previous computer experience, education level, anxiety, depression, or amount of daily texting did not show any relation to performance; however age did show a significant relationship in the opposite direction to performance on the CBDS. Results on the HADS indicated significantly higher levels of anxiety and depression in females than males. Female's performance indicated higher statistically significant levels of sustained attention (vigilance) and accuracy (prudence) than males as demonstrated by results on the IVACPT and Trails B, as well as on all CBDS items apart from the BWKPDM sequencing task. Males generally demonstrated higher levels of impulsivity as demonstrated by faster response times and greater frequency and earlier presentation of errors across the majority of CBDS tasks.

Discussion

The main aim of this study was to explore the relationship between a newly developed computerised version of the Behavioural Dyscontrol Scale (CBDS) and the original manually administered version (MBDS), and examine the extent to which the two tests measured the same construct of executive function (EF). Whilst prior research has shown the MBDS to demonstrate high levels of ecological validity when used with geriatric patients and those with TBI, the measure has enjoyed restricted use in clinical settings largely due to the reliance it has on a subjective scoring system and narrow range of scores limiting its utility particularly with younger and higher functioning populations. The BDS-EV was developed (Eastvold, et al., 2005) with the purpose of improving some of the limitations demonstrated by the original manual version. The current computerised version represents a continuation and refinement of this process to enhance the test's sensitivity to subtle executive weaknesses often undetected by traditional clinical measures of EF. The developers of the CBDS used Luria's theory of pluripotentiality to underpin design of the measure ensuring consistency with the original manual version and retaining the Three Functional Unit framework.

Findings

The first research question examined the extent to which the CBDS measured the same construct of EF as the MBDS. Creating a composite Total CBDS score for each participant was necessary in order to make this comparison. The selection of items to comprise the Total CBDS score was based on Luria's Three Functional Unit framework and previous variables used in research by Eastvold and others (2007) into the BDS-EV which analysed response times and different types of errors to assess effort, proceduralisation, attention and impulsivity. High levels of internal consistency were demonstrated using Cronbach's alpha. Results between the Three Functional Sub-units on both manual and computer versions yielded very good reliability. Following the demonstration of adequate reliability the CBDS construct validity was tested to assess the extent to which the manual and computerised versions were measuring the same executive processes. Correlation coefficients between Total scores and all Three Functional Units expressed a moderate to high degree of correspondence between CBDS and MBDS in the same direction. Therefore it could be concluded that based on this study the CBDS provides a reliable, valid measure of EF and does assess the same construct, but in greater detail (as evidenced by a broader depth and range of scores) than the manual version. This suggests that as expected scores generated by the CBDS do accurately reflect a person's strengths and weaknesses in relation to executive processes and in particular the capacity to independently plan and execute appropriate goal oriented action necessary for functional autonomy. The broad range of scores generated by participants who were from a high functioning non-clinical population suggests the CBDS has a high degree of sensitivity which in a clinical population could be extremely sensitive to subtle impairments.

The FIF, MPF, and EIF between both versions showed moderate to high correlations, with the MPF and the FIF having the most robust relationships;

however the EIF was the only sub-unit of the CBDS to demonstrate significant relationships with all measures of executive function used in this study. This finding was in contrast to previous studies (Eastvold et al., 2007) that concluded the EIF was the least robust sub-unit. Correlations between individual equivalent items on the two versions of the BDS also demonstrated strong relationships. As reported in the introduction section, previous research by Strauss and others (2006) into the manual BDS has shown administrator inexperience to be an important factor in accuracy and reliability of scores influencing test retest reliability. The high correlation coefficient shown between the individual equivalent task items was therefore relatively unexpected given the researcher was inexperienced in using and scoring the measure

In relation to Luria's theory of pluripotentiality, it is interesting to note that the Insight item on the MBDS showed a significant correlation with the Fluid Intelligence Factor sub-unit scores of the CBDS. This supports Luria's theory that working memory and insight are intrinsically related and exemplifies the multi functionality and connectedness of seemingly divergent components of EF as described by Frais and others (2006). It also supports the opinion that the CBDS can be a reliable measure of insight although the measure has no specific equivalent to item 9 of the manual version. However the recognition of insight as part of the Fluid Intelligence Factor is dependent upon the knowledge of those interpreting the scores which means it could be overlooked.

Examining convergent validity was the second research aim, and results showed this was high when equivalent tasks were compared. Both traditional clinical measures of EF used in this study, (Trails A and B, and the IVACPT) yielded significant relationships with comparable CBDS Functional sub-unit and Total Scores. The strong relationship demonstrated was between the EIF sub-unit and the FSRCQ and FSAQ (full scale response and attention control quotients) which suggests that this sub-unit is measuring impulsivity, self-regulation, and inhibition of inappropriate responses (Environmental Independence Factor). This is the sub-unit found previously by Eastvold et al., (2007) in the BDS-EV to be the least robust so using the IVACP scores to compare to CBDS scores, and finding moderate to high correlations has been useful and supports the strength and validity of this sub-unit within the CBDS. This can also be said of the Fluid Intelligence Factor sub-unit and Total CBDS scored that showed high correlations with Trails A and B. This supports previous findings by Reynolds (2002) and Strauss (2006) that the Trails correlates well with other measures of cognitive flexibility, attention and processing speed and is predictive of adaptive daily functioning capacity.

The influence of gender on task performance was the third question this study addressed. Certain features of performance were shown to characterise male and female scores, whilst overall scores for both measures were similar. Specific aspects highlighted were expected and in line with the literature discussed earlier in this study. As noted by Roebuck and others (2008) overall males were more impulsive and less accurate than females, males tended to focus more on speed, whilst females tended to focus more on accuracy. These conclusions are based on a higher number of incorrect responses being recorded by males than females on the majority of items, together with shorter length of time prior to the first error being committed. Slower response times when inhibiting a responses (as in1Tap/2Tap task) suggesting more effort is required to do this than for females was also indicated by the results, and it is interesting to reflect on how these kinds of finite details of certain aspects of performance could not have been detected using the manual BDS. Females also performed faster than males on the Alphanumeric Sequencing item of the CBDS, and on the Trails B. It is also interesting to note and further supports the convergent validity of the computerised measure, that the difference in mean sequence time for females across the ANS item and Trails B was .12 seconds, indicating the two items were closely measuring the same aspects of performance with a similar sensitivity.

With regard to other demographic variables such as computer experience, amount of daily texts, anxiety and age which have previously been shown to sometimes influence performance, only age was found in this study to demonstrate a statistically significant relationship with performance indicated by correlating in a negative direction with Total CBDS scores. This finding supports the literature by Frais and others (2006) who concluded that poor executive performance was related to increase in age. Reynolds (2002) and Paul et al., (2005) also showed that performance on the Trails test began to become slower after the age of 30 years. Higher computer experience could be expected to be associated with faster reaction times and higher level performance on sequencing tasks, however computer use was not found to correlate with performance in this study. Although the researched literature by Iverson et al., (2009) highlighted significant influence of previous computer experience on performance when using computer administered measures, this study yielded no effect for familiarity in using a computer. One possible explanation for this is that the sample comprised only tertiary students, the majority reporting a moderate to high degree of experience using a computer. Paul and others (2005) also referred to the effects of level of education on results using computer testing, this could also relate to computer experience as those with higher education are more likely to be familiar with computer technology. It seems reasonable to suggest that education level and age can both influence attitude to computer testing, and that computer specific negative affect as referred to by Albert et al., (2002) could impact on performance. Vispoel (2000) and Keith (1998) discussed participants test mode preference as a motivational factor which could negatively influence errors and response times. Literature on this by Weber et al., (2002) suggested that those with anxiety or depression preferred .pen and paper delivered tasks over computer administered tasks.

The fact that attitude and experience in using computers has been shown to influence performance on computer administered tests, raises questions about the generalisability of results from this study to a population with a more negative attitude to computers and less experience in using them.

Although the questionnaire used in this research did not specifically ask the participants about their preferences regarding interpersonal or computer administered test modes, it was noted when positive or negative comments were made about the paper and pencil tasks in comparison to the computer administered tasks. The researcher's subjective observations of this were that the participant preferences in this study were fairly evenly divided. Whilst some individuals found the computer mode less anxiety provoking than the face to face interaction with the administrator, others found the computer mode more anxiety provoking, particularly when they were unsure about when to start and stop tasks, and whether they were pressing the keys hard enough, using the correct fingers, or whether the response had been registered. It seems that performance anxiety can be increased or decreased by each mode of test delivery depending on a range of personal and contextual factors that vary to a high degree between individuals.

Gender was shown to influence performance with females exhibiting consistently more accurate and less impulsive performances than males. This was consistent with the literature (Roebuck, 2008) and suggests the measure is subject to systemic effects of different demographic variables on scores which could affect its generalisability. The manual version having a less finite specific and precise scoring system could be less vulnerable to the influence of gender on performance than the computer version. Greater sensitivity of the CBDS also brings more vulnerability to extraneous variables.

Equivalence

The scoring system of each measure, and type and depth of data extracted is very different, and one of the main areas of difference between the two measures in which equivalence is not possible. This point of difference is what sushi 2005 was referring to when he concluded that the BDS EV and in the MBDS should not be viewed as interchangeable. A limited range of potential scores was one of the main criticisms of the MBDS. Therefore the ability to record response times in milliseconds, rather than only number of errors subjectively observed by the administrator is a key feature of the computerised version. This means the measure is less likely to be subject to ceiling effects, and can measure differentiation in performance to a much finer level. Response times in milliseconds can provide detailed information about required effort and ability to proceduralise a novel behavioural sequence. For example the 1Tap/2Tap task showed response times for the two tap response became slower with time suggesting more effort is required to do this as the task continues, whereas the effort required to implement a single tap response became less with time. The extraction of the data from the CBDS is an area of technical and theoretical interpretation requiring consideration which the manual version does not have. What data is extracted influences what information can be learnt, what is accentuated, and what is omitted from a person's performance and therefore can bias how a person's capacities could be represented in results.

In contrast the BDS has a highly subjective aspect to the scoring system together with a limited range of scores possible making it prone to ceiling effects and limiting the range of information yielded. The duration of tasks is not the same on the two versions, with the CBDS tasks being longer. Elkind and others (2001) found the examiner presenting instructions could allow the participant to adaptively compensate for executive difficulties in manually administered measures. This area of weakness is addressed by using a computerised administration, however it is also important to be aware that tasks need to be time-limited and simple to execute in order to minimise participant despondency.

Mean scores in this study were high compared to previous research on the MBDS in a student population. This could be related to the researchers' inexperience in administration of the measure. However further to this, it is interesting to note that while the mean score was high on the manual version a moderate to high correlation between Total scores on the manual and computer tests in this study was still found suggesting the administrators subjective scoring was consistent with level of performance as measured by the CBDS.

Task requirements were different in some aspects (variability of cues being visual, auditory, or tactile) however results showed they were measuring the same executive processes.

This study noted comments from participants around their subjective experience of each version, and the varying responses suggest another aspect of difference between the two measures which cannot be exactly replicated because of the qualitative aspects of the task demands. For example some participants found the presence of the administrator and face to face contact in the manual version as reassuring, whereas others found this more anxiety provoking than the computer administration. Differences were also noted between participants who preferred to focus on the manipulation of the computer keyboard, rather than have to focus on interpersonal exchanges and manual hand movements. The negative influence of anxiety on performance has been noted in previous literature (Weber et al., 2007) and although in this study this was not found to have a statistically significant effect on response times, or error frequency, the participant's comments on how the two different modes of administration could reduce or increase their levels of anxiety were noted. Level of computer experience in this study was not a factor influencing performance anxiety; however in less high functioning populations lower levels of confidence and inexperience in using a computer keyboard could impact on anxiety and consequently performance.

Participants Comments

Participants' comments during and after the tasks were noted by the researcher and included the following: The computer screen flickered and a modern flatscreen would improve eye fatigue and minimise distraction.
The words red and green could be easily anticipated, using two words starting with the same letter would increase the task requirement and be more likely to produce errors in performance. When delivering this task in the manual version, the administrator could turn away from the participant to minimise the visual cue of the administrator's facial movements. The tasks providing a visual computer screen cue to end the task (BWKPDM) meant the participant was distracted from proceduralising the required response on the keyboard by having to look up at the screen and check when to stop

The sample was roughly equally divided in the preference for either the manual or computerised version of the test. Some expressed experiencing higher levels of anxiety under the face-to-face interaction conditions of the manual version, whereas others preferred this and experienced more performance anxiety when using the computer administered tasks. Several participants commented they were more motivated to try harder and get more tasks right in the manual test because it was more fun engaging with a person than with the computer.

Suggestions for improvement of the CBDS

An auditory cue to signal the start and end of the tasks would minimise distraction from the keyboard occurring on tasks that ended with a visual cue on the computer screen. This could be made consistent on all items signifying the beginning end of the task.

One suggestion for improvement of the CBDS would be to extend the length of all of the tasks in order to capture more varied depths of data and reduce ceiling effects. The Go/No-Go tasks showed the least sensitivity to variation in performance, and it could be suggested, that increasing the presentation time of items, would increase demand on attention and the likelihood of errors of commission and omission would increase as the task requirement increased. It has been noted in the literature that perseverative responses often do not emerge until the latter half of tasks, thus it is important when administering tasks to continue to do so for a long enough time period to allow for inappropriate responses to emerge. To shorten the measure, items not yielding such a range, or depth of data, could be omitted. From the results of this study, these could be the first item (R/L, L/R) as this showed no differentiation in scores from the RRL/LLR items and the ABC sequencing item, as this was only a single short sequence and, like the Red/Green item, yielded a very narrow range of scores across participants. The items yielding the most useful data across the Three Functional Units were the RRL/LLR, the Alphanumeric Sequencing and 1 Tap/2Tap tasks. An Insight item could be added if such could be designed to fit this mode of administration. However as previously mentioned, the Fluid Intelligence Factor score is recognised in the research, and shown by this study, as being related to scores that represent insight. Therefore it could be argued that the computer measure does assess insight indirectly.

Limitations of this study

The nature of our sample was limited to a tertiary student population with an average age of 25. None of this population had any difficulty with daily functioning, and all had previous computer experience. This means that the expected results of this study should have a relatively limited range of scores. As discussed earlier the fact that a wide range of scores was generated by the CBDS in this non-clinical sample suggests a high degree of sensitivity and discrimination of the measure. Although the literature on computerised testing by Iverson (2009), suggested that ethnicity was unrelated to performance, ethnicity may be related to education levels and familiarity with computers which would then influence performance. The population in this study was limited in range of ethnicity and only comprised 38 participants.

Iversen and others (2009) referred to the significant effect of previous computer experience on performance. The fact that our study did not show the same effects was probably due to the high level of education and the young age of the sample. In a 2005 study, Paul and others found that a higher age and lower education level increased performance anxiety when participants were required to use computer keyboards during tasks. Participants' reaction to mode of delivery is therefore one potential cause of error. Effect sizes for the influence of gender and scores demonstrating a relationship to age suggest the measure is subject to systemic effects of different variables on scores could affect the generalisability of the measure. These characteristics of our sample call into question the transferable validity in a population with lower educational levels and less experience in using computers.

The design of the data extraction from the digital battery software has a significant influence over results. It can shape the emphasis and range of scores that are available for interpretation.

The rich potential variety of data that can be generated by the CBDS is one significant advantage that it has over the manual version. However this advantage also means that the results and information which is then interpreted for the purposes of making useful clinical judgements are subject to the influence of the intentions, technical skills and knowledge of the designer of the specialised software. Expert clinical knowledge is therefore required to standardise and formalise methodology and interpretation. Many response characteristics could have been useful for this research, but, due to time restraints, were not used. However the scope for future research to further develop the technical aspects of data software, leading to the statistical analysis and extraction of other aspects of performance not covered in this study.

Conclusions and future directions

In terms of future research a repeated measures study would highlight test consistency and generalisability for use within different populations. In particular further testing of the CBDS could be undertaken using a mixed age non-clinical sample, a larger population with mixed educational background and ethnicities in order to extend the efficacy beyond neuropsychological rehabilitation. Another area for future research could be the influence on performance of visual or auditory learners. Use in older and younger age groups who are recognised in the literature (Blair et al.,;2012) as typically having different levels of EF development to adults, could widen the measure's applicability to non-clinical contexts, such as education and assessment of a person's capacity to self-regulate. This could be a consideration for the prediction of post injury and post illness functional outcomes. Measuring improvements in outcomes from brain injury and disease, recovery in chronic health conditions, medication adherence, as well as children's academic performance and pro-social behavioural development in the learning environment could be a useful future direction for research on the CBDS. The development of a valid computerised version of the BDS would mean more people administer it with less training.

The CBDS has shown a high degree of sensitivity to Luria’s Three Functional Units and generated a huge range and depth of data in comparison to the MBDS. The content validity looks good because we can show the CBDS items are representative of the attribute of executive function. Construct validity is high because the CBDS provided a reliable measure of set shifting, inhibition, automaticity, attention, processing speed and impulsivity which are key EF processes. The high degree of correlation between tasks on the CBDS and MBDS found in this study are to be expected because both measures were underpinned by Luria’s theory of pluripotentiality. We would expect therefore that a person’s capacity to plan and act appropriately in life, in line with their well-being and without the need for external cues, could be predicted to some degree from the scores on this test.

This study shows that the CBDS does measure the same construct as the MBDS, however it could be said from these results that the CBDS is measuring performance on a more detailed level and providing a greater range of information than the MBDS, therefore having greater clinical utility in a wider range of healthcare settings.

The question of equivalence is not clear. Even if the computer measure is shown to be measuring the same construct, based on Luria’s Three Functional Units as the manual version does, the highly specific and precise data produced remains a feature that definitively differentiates the two measures. It is important that the test scores accurately reflect what they are intended to measure, and the effects of demographic variables minimised. Different formats may be more or less useful in different clinical situations with different populations. Careful consideration of all variables is needed which will require specialised technical knowledge in collaboration with expert clinical knowledge to standardise methodology and interpretation.

The CBDS has been developed with the intention of continuing to apply Luria's Three Functional Units of Fluid Intelligence Motor Programming, and Environmental Independence, that formed the theoretical basis for the original MBDS. The recognised limitations of the manual test of subjective scoring, administrator bias, and narrow range of scores produced have all been addressed by the CBDS, which was a further refinement of the BDS EV that required specialist equipment, reducing its application practicality in a wide range of settings. The CBDS only requires the software which can easily be reproduced, and administered by a standard desktop or laptop computer.

In conclusion this study has shown that when the whole measure is deconstructed into Luria's Three Functional Units, the CBDS shows high levels of internal reliability, construct and convergent validity. It is easy to administer, score and interpret and should be of benefit to society. These are all measures of a good test. The computer administration eliminates administrator bias and generates objective, finite data providing a potentially unlimited depth and range of information about subtle aspects of a person's EF capacity, in particular, environmental independence. It delivers tasks in a highly consistent form which would be simple and practical to administer in a range of environmental and clinical settings. Whilst essentially measuring the same construct of EF as the original manual version, the results yielded in this study, indicate that it appears to measure, the construct in a broader way, generating data of a much greater range and depth than the manual version.

References

Achiron, A., Appleboim-Gavish, N., Doniger, G., Harel, Y., Lavie, M., & Simon, E. (2007). Prolonged response times characterize cognitive performance in multiple sclerosis. European Journal of Neurology, 14, 1102-1108.

Alderman, N. (1996). Central Executive Deficit and Response to Operant Conditioning Methods. *Neuropsychological Rehabilitation,* 6(3), 161-186.

Albert, A., Browndyke, J., Cohen, R., Gouvier, W., Malone, W., Paul, R., Schatz, P., & Tucker, K. (2002). Computer-Related Anxiety: Examining the Impact of Technology-Specific Affect on the Performance of a Computerized Neuropsychological Assessment Measure. *Applied Neuropsychology,* 9(4), 210-218.

Alkhadher, O., Anderson, N., & Clarke, D. (1998). Equivalence and predictive validity of paper-and-pencil and computerized adaptive formats of the Differential Aptitude Tests. Journal of Occupational and Organizational Psychology, 71, 205-217.

Ashton, V., Brooks, B., Gualtier, C., Iverson, G., & Johnson, L. (2009). Does familiarity with computers affect computerized neuropsychology test performance? *Journal of clinical and experimental neuropsychology,* 31(5), 594-604.

Bandalos, D., Garcia-Barrera, M., & Kamphaus, R. (2010). Theoretical and Statistical Derivation of a Screener for the Behavioural Assessment of Executive Functions in Children.

Barker, L., Heavey, L., McHale, S., Taylor, S. (2012). The Typical Developmental Trajectory of Social and Executive and Executive Functions in Late Adolescence and Early Adulthood. *Developmental Psychology.*

Belanger, H., Grigsby, J., Hamman, R., Malloy, P., Salloway, S., & Wilder-Willis, K. (2005). Assessing motor and cognitive regulation in AD, MCI and controls using the Behavioural Dyscontrol Scale. *Archives of Clinical Neuropsychology,* 20, 183-189.

Bender, C., Dunbar-Jacob, J., Ryan, C., Sereika, S., & Stilley, C. (2010). The Impact of Cognitive Function on Medication Management: Three Studies. *Health Psychology,* 29(1), 50-55.

Blair, C., Greenberg, M., Willoughby, M., & Wirth, R. (2012). The Measurement of Executive Function at Age 5: Psychometic Properties and Relationship to Academic Achievement. *Psychological Assessment,* 24(1), 226-239.

Bolger, J., & Suchy, Y. (1999). The Behavioural Dyscontrol Scale as a Predictor of Aggression against Self or Others in Psychological Inpatients. *The Clinical Neuropsychologist,* 13(4), 487-494.

Buxbaum, L., Dawson, A., & Linsley, D. (2012). Reliability and Validity of the Virtual Reality Lateralized Attention Test in Assessing Hemispatial Neglect in Right-Hemisphere Stroke. *Neuropsychology,* 26(4), 430-441.

Chan, R. (2001). Dysexecutive sumptoms among a non-clinical sample: A study with the use of the Dysexecutive Questionnaire. *British Journal of Psychology,* 92, 551-565.

Chervinskaya, K., & Shchelkova, O. (2005). Methodological Issues in the Development of Computerised Methods of Diagnosis in Clinical Psychology. *International Journal of Mental Health,* 33(8), 45-52.

Cope, C., Derbidge, C., & Suchy, Y. (2005). Behavioural dyscontrol scale-electronic version: first examination of reliability, validity, and incremental utility. *The Clinical Neuropsychologist,* 19, 4-26.

Diesfeldt, H. (2004). Executive functioning in psychogeriatric patients: scalability and construct validity of the Behavioural Dyscontrol Scale (BDS). *International Journal of Geriatric Psychiatry,* 19, 1065-1073.

Dixon, R., Frias, C., Strauss, E. (2006). Structure of Four Executive Functioning Tests in Healthy Older Adults. *Neuropsychology,* 20(2), 206-214.

Eastvold, A., Strassberg, D., Suchy, Y., & Whittaker, W. (2007). Validation of the Behavioural Dyscontrol Scale-Electronic Version: Sensitivity to subtle sequelae of mild traumatic brain injury. *Brain Injury,* 21(1), 69-80.

Ecklund-Johnson, E., Miller, S., & Sweet, J. (2004). Confirmatory Factor Analysis of the Behavioural Dyscontrol Scale in a Mixed Clinical Sample. *The Clinical Neuropsychologist, 18, 395-410.*

Elkind, J., Prather, P., Rosenthal, S., Rubin, E., & Skoff, B. (2001). A Simulated Reality Scenario Compared with the Computerized

Wisconsin Card Sorting Test: An Analysis of Preliminary Results. *Cyber Psychology and Behaviour,* 4(4).

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

Fritze, J., Kuhner, T., Maurer, K., Schneider, B., & Weber, B. (2002). Bias in omputerized neuropsychological assessment of depressive disorders caused by computer attitude. *Acta Psychiatr Scand,* 105, 126-130.

Giovanello, K., & Vaughan, L. (2010). Executive Function in Daily Life: Age-Related Influences of Executive Processes on Instrumental Activities of Daily Living. *Psychology and Aging,* 25(2), 343-355.

Glozman, J.M. (2007). A.R Luria and the History of Russian Neuropsychology. *Journal of the History of the Neurosciences,* 16, 168-180.

Greist, J., Kobak, K., & Reynolds, W. (1994). Computerized and Clinician Assessment of Depression and Anxiety: Respondent Evaluation and Satisfaction. *Journal of personality assessment,* 63(1), 173-180.

Grigsby, J., & Kaye, K. (1996). Behavioural Dyscontrol Scale: Manual.

Grigsby, J., Kaye, K., & Robbins, L. (1992). Reliabilities, norms and factor structure of the behavioural dyscontrol scale. *Perceptual and Motor Skills,* 74, 883-892.

Grigsby, J., Kramer, A., Myers, J., & Teel, C. (2009). Nurses' assessment of rehabilitation potential and prediction of functional status at discharge from inpatient rehabilitation. *International Journal of Rehabilitation Research,* 32, 246-266.

Groth-Marnat, G., (2003). *Handbook of Psychological Assessment* (4th ed.). Hoboken, New Jersey: John Wiley & Sons Inc.

Keith, M., Stanislav, A., & Wesnes, K. (1998). Validity of a cognitive computerized assessment system in brain-injured patients. *Brain Injury,* 12(12), 1037-1043.

Hall, J., & Harvey, M. (2008). Behavioural regulation: factor analysis and application of the Behavioural Dyscontrol Scale in dementia and mild cognitive impairment. *International Journal of Geriatric Psychiatry,* 23, 314-318.

Lam, C., Leahy, B., Suchy, Y., & Sweet, J. (2003). Behavioural Dyscontrol Scale Deficits Among Traumatic Brain Injury Patients, Part II: Comparison to Other Measures of Executive Functioning. *The Clinical Neuropsychologist,* 17(4), 492-506.

Languis, M., & Miller, D. (1992). Luria's Theory of Brain Functioning: A Model for Research in Cognitive Psychology. *Educational Psychologist,* 27(4), 493-511.

Lawrence, R, Richard, C., & Williams, L. (2005). Preliminary validity of ''integneuro'': a new computerized battery of neurocognitive tests. *Intern. J. Neuroscience,* 115, 1549-1567.

Maller, J., Ponsford, J., Rudzki, D., Spitz, G., (2012). Association Between Cognitive Performance and Functional Outcome Following Traumatic

Brain Injury: A Longitudinal Multilevel Examination. *Neuropsychology,* 26(5).

Martin, G. (2006). *Human Neuropsychology (*2nd ed.). Essex, England: Pearson Education.

Nelson, K. (2007). Psychometrics and Response Times for a Computer-Administered Mood Instrument. *Midwestern Psychological Association 2007 Annual Meeting Presentation.*

Ponsford, J., Schonberger, M., & Senathi-Raja, D. (2010). Impact of Age on Long-Term Cognitive Function After Traumatic Brain Injury. *Neuropsychology,* 24(3), 336-344.

Putz, B., & Schatz, P. (2006). Cross-Validation of Measures Used for Computer-Based Assessment of Concussion. *Applied Neuropsychology,* 13(3), 151-159.

Sherman, E., Spreen, O., & Strauss, E. (2006). A compendium of neuropsychological tests: Administration, Norms and Commentary.

Tinsley, V., & Waters, H. (1982). The Development of Verbal Control over Motor Behaviour: A Replication and Extension of Luria's Findings. *Child Development,* 53, 746-753.

Vispoel, W.P. (2000). Computerized Versus Paper-and-Pencil Assessment of Self-Concept: Score Comparability and Respondent Preferences. *Measurement and evaluation in counselling and development,* 33.

Appendix A Tables 20 – 28

Table 20 ***Descriptive Statistics for MBDS***

Measure MBDS	**Female** Mean (*SD)* N = 26	**Male** Mean (*SD)* N = 12	**Total** Mean (*SD)* N= 38	**ANOVA**	**Effect size**
RRL - finger taps	2.62 (.591)	2.33 (.778)	2.53 (.647)	F (1,36) = 1.58 p = .216	0.42
LLR- finger taps	2.46 (.582)	2.33 (.888)	2.42 (.683)	F (1,36) = 2.84 p = .598	.000
RED/GREEN go/no-go	2.69 (.618)	2.83 (.389)	2.74 (.554)	F (1,36) = .525 p = .474	0.14
I TAP / 2 TAPS	2.96 (.196)	2.92 (.289)	2.95 (.226)	F (1,36) = .317 p =.577	.009
ALTERNATE FINGER THUMB SEQUENCE	2.00 (.693)	1.83 (.718)	1.95 (.695)	F (1,36) = .465 p = .500	.013
FIST-EDGE-PALM	2.12 (.558)	1.92 (.793)	2.05 (.655)	F (1,36) = .750 P = .392	.020
HEADS test	2.42 (.578)	2.42 (.783)	2.42 (.642)	F (1,36) = .001 p = .978	.000
ALPHANUMERIC SEQUENCING	1.81 (1.02)	1.67 (.985)	1.76 (.998)	F (1,36) = .160 p = .691	.004
ALPHANUMERIC SEQUENCING-TIME	25.21 (11.03)	28.09 (1.35)	26.12 (11.06)	F (1,36) =.550 p=.463	0.15
INSIGHT	2.77 (.430)	2.58 (.669)	2.71 (.515)	F=(1,36) =1.67 p=.307	.029
TOTAL SCORE	**21.88 (2.97)**	**20.92 (4.37)**	**21.58 (3.43)**	**F= (1,36)=.645 p=.427**	**.018**

Note:
MBDS =Manual Behavioural Dyscontrol Scale

Table 21 ***Descriptive Statistics for the IVACPT***

Measure	Female Mean (SD) N = 26	Male Mean (SD) N= 12	Total Mean (SD) N= 38	ANOVA	Effect size
IVACPT					
FSRCQ	*93.38 (17.98)*	*92.75 (16.26)*	*93.18 (17.23)*	*F (1,36) = .011 P=.918*	*.000*
FSAQ	*104.0 (14.66)*	*90.08 (30.58)*	*99.66 (21.61)*	*F (1,36) = .369 P=.063*	*.093*
ARC ***Prudence***	*96.65 (15.61)*	*90.17 (20.37)*	*94.61 (17.20)*	*F (1,36) = .1.16 P = .287*	*.031*
VRC ***Prudence***	*81.27 (25.29)*	*85.33 (22.08)*	*82.55 (24.10)*	*F (1,36) = .229 P = .635*	*.006*
AA ***Vigilance***	*102.2 (86.37)*	*82.25 (34.03)*	*95.95 (21.99)*	*F (1,36)* ***= 329*** *P* ***= .007****	*.184*
AA ***Speed***	*104.8 (24.55)*	*104.0 (22.97)*	*104.61 (23.75)*	*F (1,36) = .011 P = .917*	*.000*
VA ***Vigilance***	*102.8 (9.58)*	*76.75 (40.71)*	*94.61 (26.57)*	*F (1,36)* ***= 9.80*** *P* ***= .003****	*.214*
VA ***Speed***	*104.9 (21.85)*	*95.67 30.54)*	*94.61 (24.88)*	*F (1,36) = 1.14 P = .293*	*.031*
SAAQ	*102.6 (19.14)*	*85.08 (38.25)*	*97.08 (27.40)*	*F (1,36) = 3.59 P = .066*	*.091*
SVAQ	*96.46 (28.30)*	*80.50 (45.61)*	*91.42 (34.90)*	*F (1,36) = 1.75P = 1.94*	*.046*

Table 22 ***Descriptive Statistics for R/L, CBDS***

Measure	**Female** Mean (*SD*)	**Male** Mean (*SD*)	**Total** Mean (*SD*)	**ANOVA**	**Effect size**
CBDS - Task R/L					
Total number of keys pressed	177.1 (41.80)	184.5 (63.92)	179.4 (49.06)	F (1.36) =.180 p= .674	.005
Total number of correct keys pressed	87.08 (20.46)	87.08 (32.78)	87.08 (24.54)	F (1,36) =.000 P = .999	.000
Number of keys pressed before 1st error	35.19 (43.03)	39.75 (53.14)	36.63 (46.47)	F (1,36) =.077 P = .783	.002
Total number of errors	1.31 (1.71)	3.92 (5.69)	2.13 (3.62)	F (1,36) =4.67 **P** = **.037***	.115
Fastest response	.257	.229	.248	F (1,36) =.466 P = .449	.013

Table 23 ***Descriptive Statistics for RRL on the CBDS***

Measure	**Female** Mean (SD)	**Male** Mean (SD)	**Total** Mean (SD)	**ANOVA**	**Effect size**
CBDS - Task RRL					
Total number of keys pressed	126.3 (18.87)	142.1 (55.40)	131.3 (34.70)	F (1,36) =1.75 P = 1.47	.046
Total number of correct sequences	39.92 (5.720)	41.50 (21.75)	40.42 (12.76)	F (1,36) = .122 P = .728	.003
Number of keys pressed before 1st error	29.62 (13.90)	23.75 27.02)	27.76 (18.85)	F (1,36) =.790 P = .380	.021
Total number of incorrect sequences	2.23 (2.58)	5.25 (8.08)	3.18 (5.09)	F (1,36) =3.04 **P = .040**	.078
Fastest sequence time	.610 (.124)	.595 (.240)	.605 (.166)	F (1,36) =0.58 P = .811	.002

Table 24 ***Descriptive Statistics for Red/Green Go/No-Go CBDS***

Measure	**Female** Mean (*SD*)	**Male** Mean (*SD*)	**Total** Mean (*SD*)	**ANOVA**	**Effect size**
CBDS - Task Red/Green Go/no-go					
Total number of correct responses	36.58 (.703)	36.83 (.932)	36.66 (.781)	F (1,36) =.883 P = .354	.024
Total number of incorrect responses	3.42 (.703)	3.33 (.651)	3.39 (.679)	F (1,36) =.140 P = .711	.004
Number of CORRECT RED responses	17 (.000)	17 (.853)	17 (.465)	F (1,36) =.000 P = 1.00	.000
Number of CORRECT GREEN responses	19.58 (.703)	19.83 (.389)	19.66 (.627)	F (1,36)=1.387 P = .247	.037
Number of INCORRECT RED responses	3.00 (.000)	3.17 (.577)	3.05 (.324)	F (1,36) = 2.23 P = .143	0.59
Number of INCORRECT GREEN responses	.420 (.703)	.170 (.389)	.340 (.627)	F (1,36) =1.38 P = .247	.037
Perseveration errors on GREEN command	.420 (.703)	.080 (.289)	.320 (.620)	F (1,36) =.257 P = 117	.067
Perseveration errors on RED command	1.96 (.196)	2.00 (.000)	1.97 (.162)	F (1,36) =.455 P = .504	.012
Mean red key press response time	752 (57.51)	760 (70.18)	754 (60.96)	F (1,36) =.160 P = .691	.004

Table 25 ***Descriptive Statistics for 1Tap / 2Tap CBDS***

Measure	**Female** Mean (*SD*)	**Male** Mean (*SD*)	**Total** Mean (*SD*)	**ANOVA**	**Effect size**
CBDS - Task 1 Tap / 2 Taps					
Total number of correct responses	38.46 (2.02)	37.83 (3.58)	38.26 (2.58)	F (1,36) =.478 P = .494	.013
Total incorrect responses	1.54 (2.02)	2.17 (3.58)	1.74 (2.58)	F (1,36) =.478 P = .494	.013
Correct responses before 1st error	11.69 (11.95)	11.00 (13.55)	11.47 (12.22)	F (1,36) =.026 P = .874	.001
Total correct responses to 2 tap command	19.58 (.578)	19.17 (.937)	19.45 (.724)	F (1,36) =2.76 P = .105	0.71
Total correct responses to 1 tap command	18.88 (2.05)	18.67 (3.00)	18.82 (2.40)	F (1,36) =0.66 P = .799	.002
Total incorrect responses to 2 tap command	1.00 (2.00)	1.33 (3.08)	1.11 (2.35)	F (1,36) = .160 P = .691	.004

Table 26 ***Descriptive Statistics for ABC CBDS***

Measure	**Female** Mean (*SD*)	**Male** Mean (*SD*)	**Total** Mean (*SD*)	**ANOVA**	**Effect size**
CBDS - Task ABC					
Total number of correct keys pressed	10.73 (3.29)	9.25 (3.52)	10.26 (3.40)	F (1,36) = 1.57 P = .217	.042
Total number of incorrect keys pressed	1.85 (2.50)	2.67 (1.92)	2.11 (2.34)	F (1,36) =1.00 P = .323	.027
Number of keys pressed before 1st error	9.08 (5.11)	7.08 (4.99)	8.45 (5.09)	F (1,36) =1.26 P= .268	.034
Total sequence time	7.52 (2.23)	9.59 (.374)	8.17 (2.09)	F (1,36) =**10.04*** **P = .003****	.210

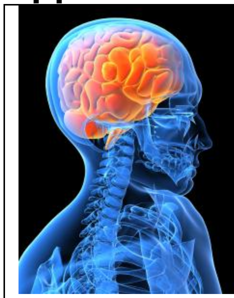
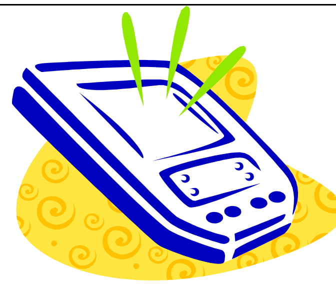
Table 27 ***Descriptive Statistics for 1A,2B,3C,4D,5E,6F,7H,8I,9J***

Measure	**Female** Mean (*SD*)	**Male** Mean (*SD*)	**Total** Mean (*SD*)	**ANOVA**	**Effect size**
CBDS - Alphanumeric sequencing					
Total number of correct sequences	2.65 (1.59)	2.08 (1.67)	2.47 (1.60)	F (1,36) =1.03 P = .315	.028
Total number of incorrect sequences	1.23 (.765)	1.33 (1.07)	1.26 (.860)	F (1,36) =.114 P= .738	.003
Total number of errors	5.38 (6.53)	9.33 (20.4)	6.63 (12.4)	F (1,36) =.826 P = .369	.022
Perseverative errors	1.00 (1.96)	.830 (1.33)	.950 (1.77)	F (1,36) =.071 P = .791	.002
Sequencing errors	4.00 (4.63)	6.83 (14.5)	4.89 (8.89)	F (1,36) =.306 P = .369	.023
Fastest sequence time	7.52 (2.23)	9.59 (.374)	8.17 (2.69)	F (1,36) = **10.0*** **P = .003****	.218
Mean Correct sequence time	10.49 (2.88)	11.91 (49.2)	10.87 (35.1)	F (1,32) =1.07 P = .307	.033
Sequence number of 1st correct sequence	1.31 (.679)	1.17 (.835)	1.26 (.724)	F (1,36)=.306 P = .584	.008
Time of 1st correct sequence	11213 (3326)	11953 (4572)	11409 (3637)	F (1,32) =.268 P = .609	.008
Time of 2nd correct sequence	9683 (2279)	10669 (4573)	9918 (2874)	F (1,19) =.436 P =.517	.022
Time of 3rd correct sequence	8963 (1423)	6656 (79.19)	8304 (1618)	F (1,5) =4.68 P =.083	.484
Digit response time	.680 (.233)	.815 (.289)	.722 (.256)	F (1,35) =.235 P =.133	.061

Table 28 ***Descriptive Statistics for BWKPDM, CBDS***

Measure	**Female** Mean (*SD*)	**Male Mean *(SD)***	**Total Mean *(SD)***	**ANNOVA**	**Effect size**
BWKPM					
Number of correct sequences	16.5 (*6.77*)	19.5 (4.58)	17.5 (6.23)	F (1,35) =1.90 P =.177	.053
Mean number of errors	.210 (.472)	.110 (.141)	.180 (.390)	F (1,35) = .446 P = .509	.013

Appendix B: Research Advertisement

	**HUMAN** **VERSUS** **COMPUTER**	

Participants wanted to validate a new computerised neuropsychological test.

This involves completing a number of human delivered and computer delivered tasks.

Your help would be much appreciated!!

If you agree to take part it will mean one session of just over an hour at the School of Psychology, University of Waikato at a time convenient to you.

To cover your expenses relating to your involvement in this project we will give you a $20 MTA voucher.

First year psychology students can gain course credits.

To be eligible you must meet the following criteria:

- • Not suffer from colour blindness
- • Have no history of head injury or psychological disorder
- Have reading and writing to at least NCEA level 1
- Speak English as your first language
- Have lived in New Zealand for 5 years or more
- Not use recreational drugs
- Be 18 yrs of age or more

To take part please contact Ruth on 0273585992, or 07 853 2952.

Email vrg3@waikato.ac.nz Or Dr Nicola Starkey, room K1.10**,** email nstarkey@waikato.ac.nz

This research has received ethics approval from the School of Psychology Ethics Committee and is supported by a University of Waikato FASS Masters scholarship

Appendix C: Information Sheet

This appendix contains the information sheet given to participants.

INFORMATION SHEET

Validation of the computerised version of the Behavioural Dyscontrol Scale

What is this study about?

When someone suffers an injury or illness, assessments are carried out to find out what impact this may have on their day to day lives when they leave hospital and return home. From this information a programme of rehabilitation is designed to help them regain the best possible quality of life. The more accurately their capabilities can be assessed, the more they and their families will be able to fully understand their potential, and what it is fair to expect them to be able to do for themselves in the long term. This study is being conducted by researchers at Waikato University and aims to find out how a computerised assessment compares to giving these tests manually. Accurate and reliable computerised assessments would mean more could be carried out, providing individuals suffering head injuries and strokes, and their families with improved access to rehabilitation that is appropriately matched to their needs.

Am I eligible to take part?

You are eligible to take part in this study if you are between 18 -70 years of age, can speak and read Level 1 English, have lived here for more than 5 years, have no history of colour blindness, head injury, or hearing problems.

What am I being asked to do?

I f you agree to take part in this study, it will involve one session of approximately 2-2.5 hours at the university of Waikato, department of Social Science at a time convenient to you. You will be asked to take part in a number of puzzle/game type tasks which most people find interesting. We will also ask for some information about you and how you are feeling. To cover your expenses relating to your involvement in this project we will give you a $20 MTA voucher. First year psychology students can gain course credits.

What will happen to my information?

Be assured that no one will be able to identify you personally. All forms will be stored in a locked cabinet, in the Department of Psychology at Waikato University. The research team will conduct the analysis of the data. At the end of the study the paper-based forms will be destroyed. We will send an electronic summary of our findings to participants who have indicated they would like to receive this information. The study has received ethical approval from the School of Psychology Ethics Committee (ethics contact person- Dr Robert Isler).

What can I expect from the researchers?

If you decide to participate in this project, the researchers will respect your right to:

- Ask any questions of the researchers about the study at any time during participation.
- Decline to answer any particular questions or carry out any of the tasks.
- Withdraw from the study at any time.

- Provide information on the understanding that it is completely confidential to the researchers. All forms are identifies by a code number, and are only seen by the researchers. It will not be possible to identify you in any articles produced from this study.
- Be given an electronic summary of findings if requested.

Who can I speak with about my participation in this project?

If you, or anyone you know is interested in taking part in this research please contact:

Ruth Kendall (masters psychology student undertaking this research) on

Phone: 07 853 2952 or mobile 0273585992 or

Email: vrg3@waikato.ac.nz

Or contact Dr Nicola Starkey on 07 856 2889 or email:

nstarkey@waikato.ac.nz

This project is supported by a University of Waikato FASS Masters scholarship award.

Appendix D: Demographic Questionnaire

This appendix contains the demographic questionnaire filled out by participants.

Demographic Questionnaire

Age

Gender

Years of education

Ethnicity

Handedness: Left Right ambidextrous (please circle one)

Occupation

Main leisure activities

- Is English your first language?

- Do you have level 1 NCEA reading or above?

- Have you lived in New Zealand for at least the past five years?

- Have you at any time suffered from any head injuries or psychological disorders?

  If yes, please give dates and brief description

- Have you any hearing or visual impairments?
  If yes, please give a brief description

- Do you suffer from colour blindness?
- Would you rate your experience with using computers as:

  LOW MODERATE HIGH (please circle one option)

- Do you use prescribed drugs?
  If yes, which ones?
- Do you use illegal drugs?
- Do you smoke?

  If yes, would you rate your smoking as SOCIAL or REGULAR (circle one)

- Do you drink alcohol?
  If yes, would you rate your consumption as OCCASIONAL
  REGULAR (circle one)
  HEAVY

- Approximate number of texts sent per day?