A comprehensive analysis of the constituents of Executive Functioning: The utility of a paediatric model towards the clear conceptualisation of EF in healthy adults

Jessica Burlak

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

College of Health and Biomedicine

Victoria University, Australia

2019

Abstract

Executive Functioning (EF) is a construct that encompasses multiple interrelated higher order skills, however, conceptualising the nebulous construct remains challenging. This paper outlined several areas of research contributing to the difficulties and challenges associated with the operationalisation of EF which pose significant challenges for the validity of psychological assessment. It was the contention of the study to confirm the validity of Anderson's (2002) paediatric model of Executive Function in a healthy adult population. Hypotheses specified that; 1) All four constructs (Attentional Control (AC), Cognitive Flexibility (CF), Information Processing (IP), Goal Setting (GS)) purported by Anderson would be upheld mathematically, 2) Attentional Control would be the strongest predictor, explaining the greatest variance in all other latent constructs. Thus, attention would be a significant domain that warrants its theoretical consideration within a model of EF, and not separate to it, and 3) Information Processing would be the second strongest predictor of other latent constructs, therefore demonstrating that IP is an influential component in a model of EF. One hundred and thirty-three adults (42 male and 91 females) aged between 18 and 50 (M=29.68, SD=7.46) completed a cognitive test battery comprising 22 tests. Of the 57 variables analysed, data reduction yielded 23 for further analyses. CFA revealed all four constructs of Anderson's model upheld, however findings imply Attentional Control as represented by more posterior attention tasks does not significantly explain EF performance at a higher level, but rather, the speed with which an individual is able to process and respond to task demands is a mediating factor in performance. The work of Peter Anderson (2002) over a decade ago has proven an exceptional platform from which to explore definitions, tests and constructs of EF, and by building his work this study has made significant advances toward a hierarchical model of EF and more importantly, the mechanisms critical to efficient functioning at the highest level of complexity. This thesis has reconciled various issues highlighted within the literature and offers a number of conclusions and directions for future research.

Doctor of Philosophy Declaration

"I, Jessica Burlak, declare that the PhD thesis entitled: A comprehensive analysis of the constituents of Executive Functioning: The utility of a paediatric model towards the clear conceptualisation of EF in healthy adults is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

Signature: [redacted] Date: 10/05/2019

Acknowledgements

To my family, David, Mum, Norm, Dad and Theresa, I absolutely have no words. You have all provided with me with the love, encouragement and support that I needed to be strong and keep going. I could not have done *any* of this without you all. I love you.

To my husband David. Where to even start. I am extremely grateful for all of your patience and support during these past four years. You have always been there for me no matter what. Your understanding, love, encouragement, and support is something that I cannot ever thank you for properly. You have a pure heart of gold because you love me unconditionally, especially towards the end of this journey when I needed you the most, and for that I will always be grateful. I am so thankful for your sacrifice over the past four years. Words will never be able to come close to express my gratitude for absolutely everything you have done for me, and the only thing I can say is that chapter 10 is dedicated to you.

To my Mum. Words cannot express how deeply grateful I am for all of your unconditional love and support for not only the past four years, but my whole life. You have given me the strength to keep pushing till the end. I am extremely thankful for the loving, nurturing and encouraging environment you have provided to me my whole life, and I could not have done any of this without you. You have always provided me with everything any daughter could ask for which has turned me into the person I am today. I know I have made you proud.

To my sister, Theresa. I cannot thank you enough for all of your support over the years, not only during my thesis, but my whole life. I am so grateful for your kind words, love and encouragement at every step of the way, and always offering to help me, and being by my side when I need it the most. You are a very calm and grounded person, and it is why I look up to you- you truly are the best sister any one could ask for.

To my supervisors, thank you for helping achieve this. To Michelle, I would like to thank you for all of your guidance and support over the past four years both professionally and academically. Your kind words, and your words of wisdom always helped me through, and I am so appreciative of your encouragement and support throughout this journey, and helping me achieve my PhD. I am also thankful for the opportunities you have presented me with, that has helped me develop my skills, grow as a person, and also excel in my career, to which I would not be where I am without your help.

To Emra, thank you for all of your guidance over the past seven years. I still look back to when you took over the role of being my supervisor during my honours year and I am so thankful for that happening. I am so grateful for you taking me under your wing because there is no way that I would be where I am today if it wasn't for you. You have guided and supported me not only in both my career and academically, but personally too. Words don't come close to expressing how appreciative I am of your dedication to me, and going above and beyond the call of duty. It is a testament to your character, for which I will always be thankful for. I am so grateful for the opportunities you have presented me with, and the others that you have gently pushed me to achieve over the years, especially being able to achieve such an accomplishment like a PhD. Never in a million years would I have ever thought I would, or even could, do this- so thank you.

Lastly, I would like to thank my PhD buddies. Being able to support one another during this time was a blessing, whether it was just being able to talk about our projects, or life in general, played a major role in our sanity! Michael and Andrew especially, thank you for all of your help, advice, and guidance in the statistics area and letting me talk your ears off about SEM! All of your advice and help was greatly appreciated, so thank you both. Lastly, thank you to all the people who participated in such a large test battery- I am so grateful that they all took the time to participate.

Table of Contents

List of Tables

List of Figures

Glossary of terms

AC- Attentional Control

ADHD- Attention Deficit Hyperactive Disorder

AGFI- Adjusted Goodness of Fit Index

AVE- Average Variance Explained

BI- Behavioural Inhibition

CF- Cognitive Flexibility

CFA- Confirmatory Factor Analysis

CFI- Comparative Fit Index

DLPFC- Dorsolateral Pre Frontal Cortex

ED- Executive Dysfunction

EF- Executive Functioning

EFs- Executive Functions

EFA- Exploratory Factor Analysis

EM- Expectation Maximisation

FL- Frontal Lobe

GFI- Goodness of Fit Index

GS- Goal Setting

IFI- Incremental Fit Index

IP- Information Processing

MAR- Missing At Random

MCAR- Missing Completely At Random

MLE- Maximum Likelihood Estimation

MS- Multiple Sclerosis

MVA- Missing Value Analysis

MVN- Multivariate Normality

NFI- Normed Fit Index

PCA- Principle Component Analysis

PFC- Prefrontal Cortex

PS- Processing Speed

RMSEA- Root Mean Square Error Approximation

RT- Reaction Time

SAC- Simple Attentional Capacity

SAS- Supervisory Attentional System

SEM- Structural Equation Modelling

SMC- Squared Multiple Correlations

SRMR- Standardised Root Mean Square Residual

TBI- Traumatic Brain Injury

WM- Working Memory

WMC- Working Memory Capacity

χ^2- Discrepancy chi Square

Glossary of abbreviated tests

AM- Austin Maze

AT-SAT- Auditory Threshold Serial Addition Test

BADS- Behavioural Assessment of Dysexecutive Syndrome

BNT- Boston Naming Task

BS-Bk- Block span backwards trials correct

BS-Fwd- Block span forwards trials correct

CNT- Contingency Naming Task

COWAT- Controlled Oral word Association Test (FAS version)

CRT-Choice Reaction Time

d2 FA- d2 test of attention commission errors

d2 H- d2 test of attention total correctly cancelled

d2 M- d2 test of attention omission errors

d2 total error- d2 test of attention omission + commission errors

d2CONC- d2 test of attention concentrate score

DEX- Dysexecutive Questionnaire

DKEFS- Delis-Kaplan Executive Function System

DS-Bk- Digit span backwards trials correct

DS-Fwd- Digit span forwards trials correct

EC- Elevator Counting

ECR- Elevator Counting with Reversal

ELF- Excluded Letter Fluency task

EPSI- Everyday Problem Solving Inventory

FSIQ- Full Scale Intellectual Quotient

GMLT- Groton Maze Learning Task

HMGT- Homophone Meaning Generation Test

PA- Picture Arrangement

PASAT- Paced Auditory Serial Addition Test

RNG- Random Number Generation

ROCFT- Rey Osterrieth Complex Figure Task

ROCFT ORG- Rey Osterrieth Complex Figure Task Organisation Score

SAT-Serial Attention Test

SRT- Simple Reaction Time

TEA- Test of Everyday Attention

TMT B-A Trail Making Test part B-A difference score

TMT-A- Trail Making Test part A

TMT-B- Trail Making Test part B

TOH- Tower of Hanoi

TOL- Tower of London

TOL-F Tower of London (Freiburg version)

TOL-R Tower of London Revised

TSC-dual task- Telephone Search While Counting Dual Task

TSC-E- Telephone Search While Counting time per target weighted for accuracy of tone counting (does not attempt to control for the decrement and individual variation in processing or psychomotor speed)

UCO- Uses for Common Objects Task

VE- Visual Elevator

VT-SAT- Visual Threshold Serial Addition Test

WAIS- Wechsler Adult Intelligence Scale

WASI- Wechsler Abbreviated Scale of Intelligence

WCST- Wisconsin Card Sorting Test

WISC- Wechsler Intelligence Scale for Children

Chapter 1 The Problems at Hand: Failures of Executive Functioning Theory

Executive Functioning (EF) is a construct that encompasses multiple interrelated higher order skills. Primarily, Executive Functions are skills that enable a person to engage successfully in complex tasks to function as an independent, socially appropriate, self-serving adult (Lezak, Howieson, & Loring, 2004). Multiple definitions exist, and certainly common to all is the consideration of EF to be a multi-faceted higher order construct in which the integrity of efficient EF is founded on intact foundation cognitive skills (Alvarez & Emory, 2006; Best & Miller, 2010; Burgess & Shallice, 1997; Lezak, 1982, 1995; Mapou & Spector, 1995; Norman & Shallice, 1986; Royall et al., 2002; Zelazo, Carter, Reznick, & Frye, 1997). As such, EF can be seen as the integration of an array of skills. Classifications vary, however generally speaking, the skills that comprise EF include the ability to problem solve, maintain and shift attention, inhibit pre potent responses, plan, implement various strategies, and utilise feedback, all of which are necessary to achieve goal-directed behaviour (Best & Miller, 2010; Royall et al., 2002). Although there is general agreement regarding the broad definitions and functions that comprise EF, there remains a considerable lack of consensus regarding their *specific* aspects and the *mechanisms of interaction* that are required to complete a novel or complex task, which pose significant challenges for the validity of psychological assessment.

1.2 The Influence of Traditional Neuroanatomy on the Conceptualisation of EF

Some of the barriers to the clear conceptualisation and development of a model of the construct of EF is that in the first instance, within the cognitive neuroscience literature, there remains a failure to offer consistent neuroanatomical findings regarding the location, or extent of excitation when an Executive Function is engaged. For example, traditional conceptualisations of the neural underpinnings of EF suggest it is mediated anteriorly. Specifically, the Pre Frontal Cortex (PFC) of the Frontal Lobes (FL) (Funahashi & Andreau, 2013; Koechlin, Corrado, Pietrini, & Grafman, 2000; Olson & Luciana, 2008; Stuss & Benson, 1984; Stuss et al., 2002; see Royall et al., 2002 for a review). With this view, a top down approach is engaged supported by a frontally mediated network that facilitates Executive Functions (Sarter, Givens, & Bruno, 2001). This provides support for the higher order conceptualisation of EF (Alvarez & Emory, 2006; Zelazo & Müller, 2002; Zillmer, Spiers, & Culbertson, 2008), particularly when a novel and complex task is engaged (Collette et al., 2005). However, current views postulate that posterior cortical regions are also necessary for efficient Executive Functioning (Banich et al., 2000; Collette & Van der Linden, 2002; Stuss & Alexander, 2000) that are responsible for selecting information based on perceptual characteristics (Mangun, 1995; O'Craven, Rosen, Kwong, Triesman, & Savoy, 1997). This view suggests a bottom up approach is engaged that triggers attentional processing utilising higher order cortical regions (Sarter et al., 2001). Thus, given that there is a proposed anterior attention system concerned with the control of attention, and a posterior attention system concerned with the allocation of spatial attention (Stuss, 2011), it has been postulated that these reciprocal projections within the PFC work both within a hierarchical yet widely distributed network (Koechlin & Summerfield, 2007). As such, questions have been raised regarding the dependency that the top down approach has on the bottom up approach, proposing the two systems interact to optimise attentional performance, where widespread cortical activation is necessary for efficient execution of EF (Egeth, & Yantis, 1997; Stuss & Alexander, 2000; Zelazo & Müller, 2002). Thus, whilst current research has progressed from traditional views of considering the frontal lobes and EF as unitary, some now debate the extent of excitation when an Executive Function is engaged given this widespread cortical activation. This is partly due to the complexity of a task that dictates the way in which information is processed in the brain (top down or bottom up). The level of complexity inherent in a task may engage different processes, therefore activate both frontal and non-frontal regions (Stuss & Alexander, 2000). To this end, given this widespread activation of different processes it is suggested that the incorporation of other cognitive domains (e.g., language, memory) is paramount for efficient execution of EF. Such considerations demand that EF no longer be considered a FL construct exclusively.

1.3 Conceptualising Specific Executive Functions

The most pervasive issue in both the adult and paediatric literature when attempting to conceptualise EF is that tests tend to be extrapolated both upwards and downwards. This poses an issue because developmentally, paediatric cognitive skills differ from those of adults. Adult level tests are not designed to track development, and paediatric tests are too simple for adults, thus denying the novelty and complexity requirements to engage the EF skill set. However, clinicians tend to use adult tests that are ‘child friendly’ and conversely, paediatric test that have been adapted to suit adults. Collectively, these limitations therefore restrict the clear conceptualisation and operationalisation of EF, particularly in adults.

As a result of the failures listed above, conceptualising the overall construct of Executive Functioning remains challenging, and assessments in particular are affected as the models of EF are essentially derived by circular reasoning. Circular reasoning implicates the diagnosis, as the tests that are used to inform a model have few psychometric properties to support it. The theory of EF is unclear because of the inconsistencies of what skills comprise it (Executive Functions). In order to remediate this, these tests have been developed from assessment paradigms in an attempt to measure and assess particular skills of EF to bring clarity to the term. However, these tests are underspecified in terms of the cognitive processes being measured and often lack validity because the theory from which these tests have originated is unclear, and as a result the cognitive skills that underpin a test are largely misattributed. Thus, in order to understand the cognitive skills that underpin a test, they are required to demonstrate higher inter correlations in order to reinforce that they represent the skills that are being measured. Essentially higher reliability coefficients are essential. Thus, whilst numerous paradigms give rise to many boutique tests that ultimately purport to measure the same thing, they have failed to do so in a psychometrically robust way in various populations as a result of the ineffective theory from which they originated.

In an attempt to bring clarity to the term, factor analytic studies have been implemented, albeit with limited success when considered collectively. Attempted fractionation of the skills that comprise EF further compounds the problem, because various factors structures have been identified in a range of populations. Thus, whilst significant gains have been made towards the fractionation of EF, the plethora of factors are likely a result of the methodological flaws of previous research. Given the limiting theory and constraints of testing, this is not an unexpected shortfall. For example, fractionation of the skills of EF was demonstrated by Miyake et al. (2000) using Confirmatory Factor Analysis (CFA) comprising a three-factor model that included shifting, updating and inhibition, which they termed 'unity and diversity' of EF. This approach helped to expand traditional views of EF away from a simple homunculus construct. Miyake and colleagues' (2000) work has guided many researchers to date, lending support to a task impurity problem where it is difficult to tease out one skill only, as a single assessment task can in fact be measuring multiple skills at once (hence the unity and diversity of their factor structure). However, their work is criticised for restricting the number of tests and preconceived factors that these might tap into (e.g., no planning, organisation) hence limiting the scope of the analysis, and validity of their conclusions. The most comprehensive approach to factor analysis was a study conducted by Testa, Bennett, and Ponsford (2012), who attempted to avoid previous methodological limitations (e.g., small sample size, limited test selection). The researchers employed a holistic approach using an Exploratory Factor Analysis (EFA) on 19 EF tests with a large sample of 200 healthy adults and found six factors to comprise EF. Although this study took a comprehensive approach, the outcome was to suggest another alternative structure, and consensus remains elusive. Furthermore, inconsistencies extend beyond the mathematical limitations that results in an excess of factors generated, but the nomenclature varies considerably. Despite this, researchers continue to employ factor analyses without due consideration of the limitations imposed by the statistical method (e.g., sample size, power analysis, type and number of tests used).

1.4 Models of Executive Functioning

Many models have been developed in order to aid the clear conceptualisation of EF and address the aforementioned issues, however some are more empirically endorsed than others creating a further barrier to model development. A model is considered an organisational framework that essentially describes a process. This review will argue that the most pervasive models that will be reviewed arguably lack formalism and universal grounding in the literature. According to Solso, MacLin, and MacLin, (2014) when the rules that underlie the model fail, they lose their "vitality as analytical or descriptive, [and should be] revised or abandoned" (p.27). Fundamental to the assessment of any cognitive skill is a sound theoretical underpinning of the classification and mechanisms associated with its function. Cognitive psychology literature has achieved clarity with respect to the definitions and underpinnings within intelligence, memory and learning, and language functioning (Adams & Gathercole, 1995; Alloway, Gathercole, & Pickering, 2006; Baddeley & Hitch, 1974), and these sound theoretical models underpin assessment, diagnosis and treatment processes. However, the cognitive construct of Executive Functioning still lacks many of the features necessary for a sound theoretical model. Executive Functioning is plagued by a lack of cohesion with respect to its definition, the skills that comprise it, and the operation and inter-relationship between these sub-skills as demarcated by models.

The foundation of cognitive psychology is based on the premise of an 'information processing' perspective (Solso et al., 2014). Since the cognitive revolution of the 1950s, theorists have concentrated efforts to understand the way in which knowledge is processed in the mind. Whilst such theorists (e.g., Miller, Bruner, Chomsky) have made significant advancements in other aspects of cognition, the domain of EF remains relatively obscured. Thus, to review that which has been stated, the collective cognitive psychology literature has failed to offer an agreed upon model of EF. EF models are typically derived from factor analytic studies that take into account outcome measures of EF tests and are performed to identify specific domains for the formation of an empirically validated model. However, despite evidence of similar factor structures across studies that potentially lend themselves to the development of a model that is generalisable to specific populations, there remains obscurity with respect to a unified model with widespread support. This is partly due to the methodological limitations of past factor analytic studies. Secondly, the tests themselves lack psychometric validity given that they have not been derived from a consistent and unified theory of EF. So the question remains; is there evidence of an empirically validated model of EF when the tests themselves lack validity?

For example, attention is a construct that underpins all conscious cognitive processing, yet there is a lack of consensus regarding the degree of overlap between attention and EF (Banich et al., 2000; Barkley, 1997; Cohen, 2014; Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991; Norman & Shallice, 1986; Posner & Peterson, 1990; Sarter et al., 2001; Stuss, Shallice, Alexander, & Picton, 1995). Whilst some have considered EF as a higher order complex cognitive construct requiring the vital skill of attention (V. Anderson, 2008) and many theorists discuss that attention plays a central role in EF (e.g., Anderson, 2002; Barkley, 1997; Posner and Peterson, 1990; Norman & Shallice, 1986), research has consistently excluded measures of attention, and the skills necessary to have efficient attention from being considered as EF in factor analytic studies, and a comprehensive inclusion of attention in models is limited. Evidence that attention must be considered when discussing EF comes from the fact that attentional difficulties manifest in many childhood disorders with an inherent executive dysfunction component (e.g., ADHD, autism, Pervasive Developmental Disorder), (e.g., Mateer & Williams, 1991) and vice versa, where behavioural implications of ADHD are associated with Executive Function deficits such as self-regulation (Barkley, 1997) planning and organisation (Grodzinsky & Diamond, 1992). Furthermore, neuroimaging studies consistently implicate the anterior and prefrontal regions of the brain when attention tasks increase in complexity, or when certain elements of attention are engaged (Banich et al., 2009; Fuster, 2002; Stuss et al., 1995; Stuss, Toth, Franchi, Alexander, Tipper, and Craik, 1999; Stuss, 2006) which are the same brain regions implicated in the mediation of EF performance. For example, shifting, selective, and divided attention, as well as inhibition, have been proposed to be mediated by the anterior regions of the brain (Stuss et al., 1995; Stuss et al., 1999; Stuss, 2006; as reported in Lezak et al., 2004, p. 80).

Currently, inconsistent and poorly endorsed theories of EF circulate within both the research and clinical paradigms (e.g., Executive Function: P. Anderson, 2002; Behavioural Inhibition: Barkley, 1997; Four component conceptualisation: Lezak, 1982; Lezak, et al., 2004; Supervisory Attentional System (SAS); Norman & Shallice, 1986; Shalllice & Burgess, 1996; Stuss et al., 1995; Problem-solving framework: Zelazo, et al., 1997). These models share similarities in that they all propose multifactorial skills to comprise EF. For example, Barkley (1997), Lezak and colleagues (2004) and Zelazo and colleagues (1997) emphasise a hierarchical placement of skills in overall EF processing. However, they differ in their putative descriptions of such skills, and how they operate to achieve efficient goal-directed behaviour. For example, Barkley’s model proposes Behavioural Inhibition (BI) is essential for EFs to occur, essentially describing inhibition as a pre-requisite skill, where

Zelazo and colleagues critique the way in which basic processes in isolation such as inhibition fail to provide enough detail regarding the inter-operationalisation of the complex nature of EF. Furthermore, whilst Stuss and colleagues' (1995) adaptation of Norman and Shallice's SAS model actually accounts for, and includes higher level attentional processes as opposed to foundational attentional skills in the role of EF, their model fails to consider how an outcome is reached, which is something that Zelazo and colleagues' model offers. The conceptualisation proposed by Lezak (1982) which was later updated with her colleagues (2004) could be argued to be the most guided definition of EF whereby "Executive Functions consist of those capacities that enable a person to engage successfully in independent purposive self-serving behaviour" (p.35). Though more of a framework than model, Lezak's four main aspects of Executive Function include volition, planning, purposive action and effective performance. Within these distinct categories, there is a set of "activity related behaviours" (p.611). Whilst these authors acknowledge the multiplicity of the skills that fit within the overarching term of 'Executive Functioning', there is a lack of clarity regarding what these skills actually are, the interrelationships between them, and how they are assessed. Whilst they have made significant gains towards identifying the skills or capacities that fall under these categories and proposing a sequence of stages (hierarchy) of cognitive processing necessary for the overall performance of any complex task, there remains a failure to consider an attention component in the overall model. A further limitation is the lack of bidirectional relationship between categories, considering there is little disagreement within the literature that these skills relay critical information back and forward between each other (Anderson, 2002; P. Anderson, 2008).

The obscurity of this construct undermines the assessment, identification and treatment of executive dysfunction, particularly in adults. A variety of EF tests are used to inform the models of EF and possible composite skills that underpin it, however the flaw of

EF both as a theory and in its assessment is that clinicians tend to draw conclusions inefficiently, due to the lack of clarity regarding how these skills are delineated. Thus, contradictory models of EF exist, which has hindered the progress towards a psychometrically robust model of EF. Thus, a review of the EF literature implores that any model of EF must consider a) separate subdomains of EF skills, b) that subdomains be considered hierarchical and c) the directionality and relationship between each of the EF skills is established.

1.5 The Utility of a Paediatric Model of Executive Functioning for Healthy Adults: Anderson's Model of Executive Function

P. Anderson's (2002) conceptually driven model (Figure 1 below) has been proposed in an attempt to describe a potential model of EF in children. This impressive model incorporates developmental, neuropsychological, and factor analytic literature that addresses many of the aforementioned limitations of describing, defining, and conceptualising EF, and has been used as a guide empirically with paediatric populations (Bodimeade, Whittingham, Lloyd, & Boyd, 2013; Høie, Mykletun, Waaler, Skeidsvoll, & Sommerfelt, 2006; Long et al., 2011). Anderson's model proposes there are four distinct sub domains supported by frontal systems that interact in an integrative manner to comprise an overall control system (Anderson, 2001). The clear strengths of this model include the fact that there is a clear attention component, as well as the existence of the bidirectional relationships between sub domains. However, what is missing from this model (but was addressed by Lezak in adults) is an explicit hierarchy of skills (i.e., there is no start and finish in the overall efficient performance of a complex task).

Anderson's model appears to encompass many of the necessary constituents regarding Executive Functions and pays homage to a range of theoretical propositions. For example, this model has demonstrated a clear integration and overlap with respect to many sub domains. This could suggest and address the contention in the literature regarding the consideration of attention as an Executive Function (executive attention) that is anteriorly related. Anderson incorporates into this domain selective attention, self-regulation, self-monitoring and inhibition, thus aligning closely with Barkley's contention that inhibition is a requirement to attentional capacities. Further, the bi-directionality of the domains in the Anderson model appear to represent the fundamental mechanisms of EF, which many other models have neglected. This is indicated by the depiction of arrows of which some demonstrate bi-directionality, and also acknowledge unidirectional pathways that feed information to other systems. Moreover, whilst Anderson's model does not have an explicit conceptualisation of a hierarchy of skills akin to Lezak's four component conceptualisation, it does however, seem to include detail regarding foundation skills necessary for the effective performance of higher order skills, equally as important is the consideration of bi-directionality.

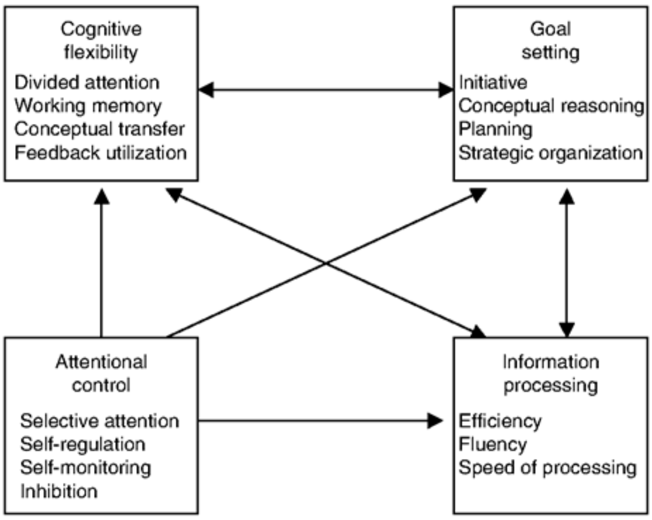


Figure 1. Anderson's (2002) proposed model of Executive Function. Reprinted from "Assessment and Development of Executive Function (EF) During Childhood" by P.

Anderson, 2002, *Child Neuropsychology, 8(2),* p.73. Reprinted with permission from personal communication with P. Anderson, April 9, 2019.

1.5.1 Rationale, contribution and significance

Executive Functioning tasks inform the models, however beyond the consensus that there is a diverse set of skills that comprise EF, there is controversy regarding a number of features and therefore, operationalising EF remains challenging. In addition, given the lack of consensus regarding the skills that comprise EF, there remains obscurity regarding a) how these skills contribute to the overall achievement of "independent purposive self-serving behaviour" b) which skills in the hierarchical chain of cognitive processes contribute more towards the overall performance of EF, and c) the degree to which the directionality of these skills interrelate. Therefore, it is the aim of the current study to confirm the validity of Anderson's (2002) paediatric model of Executive Function in a healthy adult population. This is in careful consideration of the contentious literature surrounding EF, in particular the approaches used to bring clarity to Executive Functioning.

For a variety of reasons that will be reviewed, EF theory, assessment and diagnosis of dysfunction is plagued by the lack of formalism inherent in the models used to guide research and clinical practice (Solso et al., 2014). This paper will outline several areas of research contributing to the difficulties and challenges associated with the operationalisation of Executive Functioning and will also address the problems with attribution of the constituents of EF. Contemporary models will also be presented and critiqued in light of these issues, which will highlight the significant role of Anderson's model within the EF literature.

The significance of the present study has a dovetailed approach. It will contribute toward our understanding of the differentiation of skills classified as executive. Doing so will assist future researchers to develop a psychometrically robust battery of EF that will promote scientific validation and replication of EF within an adult population using Anderson's model. Based on the findings from Anderson's validated model in an adult population, it will also have significant clinical implications for the diagnosis and management of a range of adult disorders with an inherent executive dysfunction component.

Chapter 2 Executive Functioning: Challenges in Conceptualising a Nebulous Construct

Theorists have concentrated their efforts to understand the construct of EF, and whilst consensus has been reached regarding the general applicability of the term, there remains obscurity with the constituents that underlie the overarching label. Literature consistently indicates that EF is a *set* of higher order cognitive processes necessary for goal-directed behaviour that coordinate, conduct, and supervise complex cognitive performance (Best & Miller, 2010; Hughes, 2013; Jurado & Rosselli, 2007; Lezak et al., 2004; Olson & Luciana, 2008; Royall et al., 2002; Stuss & Alexander, 2000). Others suggest EF allows goals and actions to take place (Luria, 1973), and that the set is necessary to carry out efficient complex or 'higher order' cognitive tasks, concentrating efforts for goal-directed behaviour. EF has been referred to as the highest level of human functioning (David, 1992, as cited in Anderson, Jacobs, & Anderson, 2008), as EF is not exclusive to just cognitive activity as emotional and social behavioural regulation may also play a role (Gioia, Isquith, Guy, & Kenworthy, 2000; Gyurak, Goodkind, Kramer, Miller, & Levenson, 2011; Lezak et al., 2004; Norman, & Shallice, 1986; Royall et al., 2002; Zillmer et al., 2008). To this end, little clarity exists beyond this consensus.

2.1 Neurological Underpinnings of Executive Functioning

Traditionally, research has demonstrated Executive Functioning as mediated by the anterior regions of the brain (frontal regions) (Anderson, Levin, & Jacobs, 2002; Koechlin et al., 2000; Stuss & Alexander 2000). Primarily, EF is subsumed in the Pre-Frontal Cortex (PFC) (Funahashi & Andreau, 2013; Koechlin et al., 2000; Olson & Luciana, 2008; Stuss & Benson, 1984; Stuss et al., 2002; see Royall et al., 2002 for a review) and Dorso-Lateral Pre Frontal Cortex (DLPFC) (Grattan & Eslinger, 1991; Zillmer et al., 2008), where it is thought to be the 'on-line' system that holds information (Goldman-Rakic, 1996). The PFC generally, and more specifically the DLPFC, is assumed to be responsible for most complex cognitive abilities (Diamond, 2000; Zillmer et al., 2008). Developmentally, the frontal lobe and its subsumed cortices are last to develop and are rudimentary in the early stages, where they undergo refinement and maturation as the developmental trajectory increases. The prefrontal regions continue to develop into early adulthood (Fuster, 2002; Romine & Reynolds, 2005) where neuronal proliferation and differentiation, axonal and dendritic arborisation, synaptogenesis, synaptic pruning and myelination occur (Fuster, 2002; Jurado & Rosselli, 2007; Kuan, Roth, Flavell, & Rakic, 2000). The greater density within the PFC compared to other cortical areas (Elston, Benavides-Piccione, & DeFelipe, 2001) is attributed to the myelination process that facilitates rapid transmission of electrical signals (Fuster, 2002) necessary for the integration of more complex material (De Luca & Leventer, 2008; Fuster, 2002).

However, issues arise when attempting to conceptualise this. This is partly due to numerous studies assuming frontal lobe functioning to be equivalent to Executive Functioning, because patients with frontal deficits score poorly of EF tasks. This ultimately led to the terms 'Executive Functioning' and 'frontal lobe functioning' to be used synonymously in the literature. This ambiguity between the two has led to the parallel fashion of devising neuropsychological 'frontal lobe measures' to assess Executive Dysfunction (ED). This premise asserts that those who score poorly on a range of EF tests have a frontal lobe lesion. Likewise, those who have frontal lobe lesions are hypothesised to score poorly on a range of EF tasks. This contention, however, does not always hold because EF does not immediately imply the frontal lobes and frontal lobes only, because patterns of performance are not uniform across EF tasks suggesting the existence of multiple EF skills (Packwood, Hodgetts, & Tremblay, 2011). In their meta-analytic review study, Alvarez and Emory (2006) found that some patients still perform within normal limits on the Executive Functioning measures despite having a frontal lobe lesion. Similarly, the authors also found that some people who score poorly on EF tests have no recognized frontal lobe lesions (Alvarez & Emory, 2006; Anderson, Damasio, Jones, & Tranel, 1991). Therefore, whilst many EFs may be subsumed by the PFC, other cortical and subcortical regions can be implicated in EF and performance can be affected for other reasons than the FL (Stuss & Alexander, 2000). For example, orbital regions of the PFC may also be associated with EF, given that emotional regulation may also play a role in EFs (Zelazo & Müller, 2002). Furthermore, in Alvarez and Emory's (2006) meta analytic review, they found a variety of studies highlighting the Wisconsin Card Sorting Test (WCST) - a typical measure of EF, to activate other regions than the frontal lobes such as the basil ganglia, inferior parietal cortex, occipito-temporal pole, temporo-parietal association cortex, and occipital cortices. Current views also suggest the role of posterior cortical regions (Banich et al., 2000; Collette & Van der Linden, 2002; Stuss & Alexander, 2000). Thus, current views suggest EF and frontal lobe pathology should not be used synonymously as the engagement of EF is not restricted to frontal lobes, and the integrity of the entire brain is necessary for efficient EF (Anderson, 1998; Della Salla, Gray, Spinnler, & Trivelli, 1998; Stuss & Alexander, 2000 Zelazo & Müller, 2002), as global dysfunction is rare (Zillmer et al., 2008).

Whilst significant advancements towards the structural imaging techniques used to inform the neuroanatomical location to isolate specific neural substrates associated with EF have made gains, it is unfortunate that the functional imaging techniques fall short. Imaging tells us what parts of the brain do what (because of neurological activation during task performance), however they do not necessarily tell us the how the brain works as a system, thereby inferring structure more than function, often lacking the residual deficits (Lezak et al., 2004). This may be why cognitive assessments are inconsistent with patterns of neuroanatomical findings (Snowden, 1997, as cited in Lezak et al., 2004). It is for this reason that the cortical attribution of EF is limited due to paucity of specificity. Whilst it is not appropriate to conclude a patient's FL functioning based purely on their EF performance alone and vice versa, the frontal lobes have guided researchers towards the understanding of EF, because top down modulation for complex higher order thought is necessary for efficient Executive Functioning, and the integration of information from the posterior cortices would not be possible without the frontal lobes.

One reason to possibly explain the widespread cortical activation that can occur during EF performance is by conceptualising cognitive load. Executive Functioning, by definition, is engaged when a task is novel and complex (P. Anderson, 2008) which gives rise to EF being associated with the FL. However, the perceived level of complexity inherent in a task differs amongst individuals which in turn influences the way in which the cognitive load is processed in the brain as a whole.

2.2 Engaging the Executive Functioning Skill Set

Without argument, researchers have consistently indicated that EF is engaged when the task demands are novel and/or complex, particularly when automatic tasks require additional control (Norman & Shallice, 1986; Stuss & Alexander, 2000). Whilst it is acknowledged that there is limited cortical attribution of EF due to paucity of specificity, traditional imaging studies provide support for the complex cognitive activities being mediated by the DLPFC and PFC that are subsumed in the FL (Diamond, 2000; Duke & Kaszniak, 2000; Grattan & Eslinger, 1991; Stuss et al., 2000). In particular, the lateral PFC

(Banich, 2009; Collette et al., 2005; Wager & Smith, 2003) and medial PFC (Derrfuss, Brass, Neumann, & von Cramon, 2005). These prefrontal regions support the notion of a top down approach as an increase in activation is evident when a new task is engaged (Duncan & Owen 2000, Poldrack et al., 2005).

When a top down approach is engaged, cognitive performance elicits slowed response times as higher levels of cognitive control are recruited (Shiffrin & Schneider, 1984). In contrast, when a bottom up approach is engaged this elicits a quicker response time because of the automaticity of the task where less cognitive control is required (Shiffrin & Schneider, 1984). This approach is supported by limited activation of the lateral PFC when the novelty has faded (Chein & Schneider 2005; Landau, Garavan, Schumacher, & D' Esposito, 2007). In essence, EF plays a significant role in tasks that are fluid in nature (Purdy, 2016) that require novel problem solving with the application of new learning, or the re-application of previous knowledge in a unique way. This is because while a simple task can only be considered novel once, a complex task may remain difficult with novelty that has diminished until it is mastered. Evidence supporting this notion comes from research that demonstrates EF can also be activated during well-learned behaviours, as some well-learned routinised tasks can also be consequentially complex (Stuss & Alexander, 2000).

2.2.1 Processing capacity

Cognition is a finite resource. The way information is processed depends in part upon the cognitive load required to complete it. More complex tasks necessarily require more of this finite resource, whereas simple tasks are completed very quickly, and with more automaticity, thereby using less resources. However, issues that are problematic with the proposition of EF mostly being engaged when a task is complex, is that the level of executive control required for a 'novel' or 'complex' task differs between individuals and depends on the task demands (Stuss & Alexander, 2000). Ultimately, processing capacity is influenced by the cognitive load (simple or complex) required to complete a task. This varies depending on what one considers to be a simple or complex task, thereby contributing to the inconsistencies of demarcating the neuroanatomical extent or excitation of EF.

Several theories lend themselves to explain perspectives on how cognitive load is processed. Execution of dual tasks in particular provide insight into cognitive load and processing capacity where salient information is processed serially, in parallel, or sometimes both. The Bottleneck theory of attention proposes that the human cognitive system is limited in its information processing capacity, so when a dual task is employed where the demands of a task require one to carry out two tasks simultaneously, or in close succession, only one can ‘access’ or act on this input at a time often resulting performance impairment (Lehle, Steinhauser, & Hübner, 2009; Pashler, 1994a; Tombu & Jolicoeur, 2003). Whilst research suggests that two tasks compete for this limited capacity ultimately leading to serial processing (Pashler, 1994b) typically when control demands are higher (Luria & Meiran, 2005), recently, research has also suggested that this limited processing capacity can in fact be shared, ultimately leading to parallel processing (Lehle, et al., 2009; Tombu & Jolicoeur, 2003), when levels of control are lower (Luria & Meiran, 2005). Different strategies may be employed consciously or unconsciously to accommodate the two tasks that are competing for the limited capacity information processing system, as access to both concurrently is difficult (Pashler, 1994a; Lehle et al., 2009), and questions remain as to which strategy is considered most ‘optimal’. However, theorists have proposed that there may be a choice between strategies (Lehle, et al., 2009), that may be influenced by a variety of factors, such as the type of content, sensory inputs, produced responses and internalised thoughts (e.g., cross talk model) (Pashler, 1994a). However, trade-offs between goals and strategies as well as mental effort used to obtain these goals are also a necessary consideration (Hockey, 1997). Lehle et al. (2009) investigated this theory and measured mental effort via heart rate and skin conductance (purported measures of mental effort). They found that parallel processing was associated with performance costs compared to serial processing, and serial processing was considered more effortful.

Thus, whilst several competing theories have been proposed regarding cognitive processing capacity, “it is unknown whether lower-level decisions start only when higher-level decisions have completed or vice versa (a serial model), or whether various levels of the decision are processed simultaneously, with higher levels continuously constraining lower levels (a parallel model)” (Ranti, Chatham, & Badre, 2015, p. 206). It is likely a result of the individual strategies that can be employed to complete a task (Lehle et al., 2009; Ranti et al., 2015) which are likely mediated by the cognitive load in the perceived level of complexity. Thus, as implied in the definition of EF being engaged when a set of higher order skills are required for complex decision making, conceptualising the term is further compounded when different strategies are employed to complete a task, especially when the subjective experience is difficult to discern.

2.2.2 Trade-off between speed and accuracy

For purposes of psychological assessment, satisfactory performance on a task is determined by either the speed in which a task is completed, or the accuracy of completion (e.g., error performance). Some current assessments of EF incorporate both speed and accuracy as outcome measures, however both are not always taken into consideration in the final outcome of results relative to normative data due to traditional scoring protocols. It is vital that both domains are considered equally, and the degree to which one is more influential than the other for efficient performance on an EF task, as EF cannot be measured without the consideration of the two (Anderson, 2002). This is because EF is closely related to processing speed because speed is a basic cognitive function that mediates higher cognitive processes (Pires et al., 2018), where a bidirectional relationship may be evident (Cepeda, Blackwell, & Munakata, 2013).

Researchers have explored processing speed as being directly related to accuracy on a range of cognitive tasks as poor performance in one aspect typically influence how performance is exhibited in the other. In essence, errors take time, and the less errors one makes, the faster one's speed should be. However, this premise does not always hold, and it is important to consider the effect of trading the speed one completes a task to reduce the amount of errors (e.g., taking longer to reduce errors), or trading the amount of errors to finish more quickly (e.g., making lots of mistakes to finish more quickly), and the implication on performance. For example, research has consistently demonstrated that EF suffers when the speed at which one completed a task is affected. As explained above, when a top down approach is engaged, cognitive performance elicits slower response times as this recruits higher levels of cognitive control (Shiffrin & Schneider, 1984). In contrast, when a bottom up approach is engaged performance is typically faster because of the automaticity of the task where less cognitive control is required (Shiffrin & Schneider, 1984). Thus, it is not surprising that processing speed subserves many higher order cognitive domains, particularly when processing speed is affected when tasks are complex and require more effort and cognitive resources in a range of clinical and healthy populations (Arnett et al., 1999; De Sonneville et al., 2002; Diamond, Johnson, Kaufman, & Graves, 2008; Donkin, Little, & Houpt & 2014). These trade-offs have been investigated in a range of clinical and healthy populations, however the reasons as to why one might trade remains unclear. For the most part, a trade-off is most likely due to cognitive impairments in clinical populations, however for healthy adults this is not so clear. Research has demonstrated that we are able to flexibly adopt a speed-accuracy trade-off or accuracy-speed trade-off depending on the task demands, instructions (Howell & Kreidler, 1963), payoffs (Swensson, 1972), deadlines/ time pressure (Pachella, Fisher, & Karsh, 1968) and motivation (Higgins, 1997). This has been evident from previous studies in relation to clinical groups with poor EF performance. For example, Donkin et al. (2014) found slower responses tend to be more accurate than faster responses in a sample of 8 adults, where workload capacity influenced the type of decision making, depending whether participant were in a speed or accuracy condition. Similarly, Pachella et al. (1968) found a decrease in accuracy of responses when subjects were forced to quickly make judgements, and lastly, Howell and Kreidler (1963) found there are critical speed and accuracy levels with which an individual is unwilling to perform. The authors found during an accuracy emphasis trial, participants increased their accuracy criterion, however little reduction in response speed was observed. Furthermore, during a speed emphasis trial participants demonstrated little increase in response speed, however a very large decrease in the accuracy criterion. These findings suggest that participants were unwilling to lower their accuracy criterion despite the requirement to slow down to increase their accuracy.

These findings perhaps could be explained by serial or parallel processing. Given that we are free to choose between different strategies this premise may explain why there is such a discrepancy between trading speed for accuracy or vice-versa. It may depend on the perceived complexity levels of the task, the amount of mental effort, or the cognitive capacity allotment. The consideration of the trade-offs, goals, strategies, and mental effort as described above in relation to serial or parallel processing, comes from the distinction between effectiveness and efficiency that should be considered when analysing task performance. Eason (1963, as cited in Kahneman, 1973) states that "effectiveness is a measure of the quality of performance, while efficiency is the relation between the quality of performance and the effort invested in it'' (Kahneman, 1973, p. 181).

Processing speed has been investigated in not only EF but a range of clinical populations. For example, Multiple Sclerosis (MS) (Drew, Starkey, & Isler, 2009), TBI (Campanella, Skrap, & Vallesi, 2016), reading in children (Peter, Matsushita, & Raskind, 2011), Working Memory (WM) in children (Fry & Hale, 2000) and normal ageing (Cerella, 1985; Fisk & Sharp 2004; Salthouse, 1996), where a decline in speed as age increases has significant effects on cognitive functioning (Salthouse, 1996). Furthermore, standard intelligence testing includes a Processing Speed Index is used to measure non-verbal mental and psychomotor speed (Wechsler, 1997b), where processing speed is influential to other cognitive skills. For example, processing speed can be used as a direct measure of attention, because slowed processing speed is typically an indication of attentional problems (Lezak et al., 2004). However, the most prominent issue when attempting to understand processing speed is that it is measured in a variety of ways. For example, the use of attention tasks, complex tasks, or even simple Reaction Time (RT) tasks, and as such, little consistency exists which undermines conclusions drawn. For example, processing speed has been used interchangeably with information processing speed, complex attention, reaction time, and cognitive speed, all of which are measured with different tests and task demands (Cepeda et al., 2013; Chiaravalloti, Christodoulou, Demaree, & DeLuca, 2003). A broad definition proposes that “processing speed is underlying cognitive efficiency at understanding and acting upon external stimuli, which includes integrating low level perceptual, higher level cognitive, and output speed” (Shanahan et al., 2006, p. 586). As a result, the construct validity of processing speed remains elusive (Cepeda et al., 2013), where some have suggested to remove the influence of processing speed to understand the contributions of other higher order cognitive processes (Cepeda, et al., 2013). This is most likely why most processing speed measures are simple tasks, so that the contribution of other higher cognitive functions is minimised (Fry & Hale, 2000). This ambiguity in this construct leads researchers to question what the role of processing speed is especially to EF, and if this domain specific or a more globalised function.

Domain specific speed of processing refers to a specific domain or construct comprised of different processes that may develop at a unique rate, yet vary across domains (Kail & Miller, 2006). For example, Pires et al. (2018) found a Processing Speed factor in healthy young adults highlighting that EF and processing speed are related but separable. Similarly, Kail and Miller (2006) found domain specificity for processing on language tasks compared to non-language tasks. Although not a factor analytic study demonstrating support for a construct, Drew et al. (2009) explored the relationships and the predictive validity between individual tasks and found that information processing speed is not a unitary construct in MS patients. Interestingly, Chiaravalloti et al. (2003) also found a non-unitary construct in a sample of mixed subjects (N=92), however found a difference between simple and complex processing speed. Their study used a variety of Simple Reaction Time (SRT), Choice Reaction Time (CRT), and Serial Attention Test (SAT) measures. Their factor labelled simple Reaction Time comprised the auditory SRT, visual SRT, auditory CRT, and visual CRT, where participants were required to press a button as soon as they heard or saw (or required to make a choice between two in the CRT tasks) the stimulus respectively. Their factor labelled as complex Information Processing comprised the Paced Auditory Serial Addition Test (PASAT; participants are required to add digits to a series of strings immediately after is it verbally presented; 4 different trials varying in speed of presentation), the Auditory Threshold Serial Addition Test (AT-SAT; index of speed of information processing controlling for accuracy), and the Visual Threshold Serial Addition Test (VT-SAT; identical to the AT-SAT). Furthermore, both visual and verbal measures relating to new learning abilities also loaded onto the complex Processing Speed factor; the Selective Reminding Test and 7/24 trials to criterion.

Conversely, global processing speed has also been proposed. Global processing refers to a systemic mechanism that is not specific to a particular task (Kail & Miller, 2006). This has been demonstrated by a unified processing speed factor that varies across a variety of tasks, particularly in dyslexic children (Peter et al., 2011), where white matter volume that supports axonal speed of rapid electrical firing is thought to play a role (Peter et al., 2011). This global mechanism has been supported by a range of studies, in particular global developmental trends that increase and decrease with age, and that speed measures are equally related to similar measures (Fry & Hale, 1996; Hale & Jansen, 1994; Peter et al., 2011).

Therefore, investigating the approach to an EF task is intricate particularly when the subjectivity of a complex task influences the way in which cognitive load is processed, and whether skills are employed on a conscious level, automatic level, or perhaps masking an impairment. Thus, it is vital to consider both facets of speed and accuracy in assessments of EF because they can provide an understanding of the multiple processes required when competing an EF task, which may elucidate a paramount impairment, or simply highlight a strategic performance.

2.3 Placement of Executive Functioning within an Overarching Framework

Given the widespread cortical activation that is associated with EF performance and the limited clarity regarding how the *set* of skills to comprise EF relate to one another, many theorists have proposed there may be a hierarchy *within* the set of skills of EF, but also if there is a hierarchy *between* EF and other cognitive domains, which may help explain the varying degrees of task performance (Goldstein & Green, 1995). Thus, the overall framework for cognitive assessment remains challenging, especially when it is difficult to place EF relative to other domains.

For example, a person with damage to the occipital, temporal, subcortical or parietal regions may display an array of impairments, however may also fail ‘frontal lobe tests’ (Anderson et al., 1991). Similarly, language problems may arise from self-regulatory and organisation deficits associated with Executive Functioning (Stuss & Benson, 1990 as cited in Lezak et al., 2004). Furthermore, memory is a construct that is central to many other modalities that can be affected by information processing speed, concentration, effort, organisation strategy, and self-monitoring (Ganor-Stern, Seamon, & Carrasco, 1998; Howieson & Lezak, 2002 as cited in in Lezak et al., 2004). Thus, it is not surprising when patients complain of memory and learning problems that it may in fact be secondary to their inability to attend to stimuli, which is why attention is the first construct typically examined (Lezak et al., 2004; Mirsky, Fantie, & Tatman, 1995).

As previously stated, in a latent sense there is evidence of a hierarchy of skills not only within the set of EF skills, but also between EF and other cognitive domains. Researchers have suggested that higher order cognitive functions suffer if there is damage to the human central nervous system, and any failure or disruptions in this system will most likely result in reduced output (Mirsky et al., 1995). The most pervasive issue when attempting to understand why one may fail a test, is that it remains unclear where a disruption in the overall chain of processing skills occur, which skills are proposed as being lower or higher order, and the impact this may have on other abilities. Higher order mental functions refer to complex self-regulated, mediated, and learned mental functions (Luria, 1980; Vygotsky, 1978), where these higher functions “integrate others that are considered to be basic such as attention, perceptions and memory” (Baron, 2004, p.134). Similarly, Stuss and Benson (1986, as cited in Mapou & Spector, 1995) proposed tiers of abilities. For example, abilities at the lower tier include attention, memory and language, and those at the higher level of behaviours include goal selection, planning and monitoring. However, when assessing planning and monitoring, it is difficult to tease out the skills subsumed to identify if a deficit is due to one’s planning skill (direct effect), or below in the chain of processing where an attentional or memory deficit may be the primary cause (indirect effect), and the planning deficit is the secondary impairment. Several attempts to distinguish between lower and higher level abilities have been proposed, where the removal of higher level abilities such as anticipation and planning may reveal true impairments on lower level abilities such as visuomotor abilities (Luria, 1973). Similarly, qualitative observations have also been noted as being critical to clinical assessments (Lezak, 1982). Demarcation of these skills therefore is important from a theoretical viewpoint, and critical from a clinical viewpoint, where identification of a primary vs secondary impairment is essential if remediation of impairment is to occur. It is therefore imperative that a thorough clinical assessment of lower order functions is carried out prior to investigating an EF deficit.

2.3.1 Mapou's framework for cognitive assessment (1995)

Mapou (1995) has attempted to classify a range of integrated modalities as a framework for the assessment of cognition overall. The order of placement is not on the basis of cognitive complexity, but rather to infer that skills at the lower level "are seen as *fundamental* to effective expression of remaining skills in the framework" (Mapou, 1995, p. 299). This framework comprises *global functioning, foundation skills, modality specific skills and integrated skills,* as outlined in Figure 2.

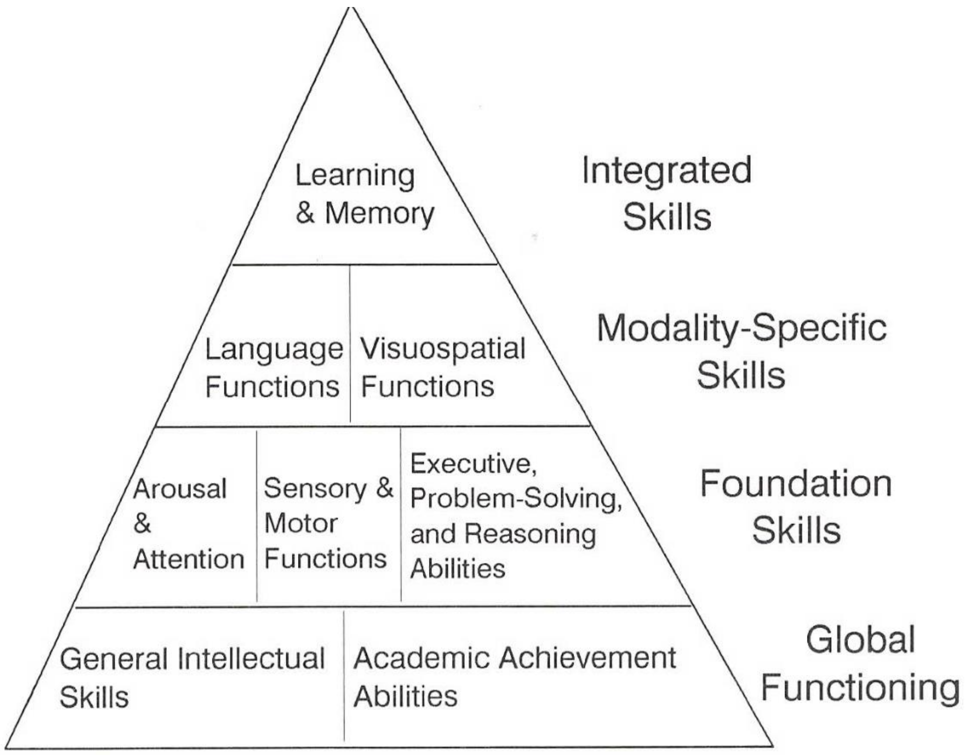


Figure 2. Schematic diagram of the framework for cognitive assessment (Mapou, 1995).

Contention exists, however, in relation to the placement of EF. The placement of EF is seen as in line with foundation level skills even though it has previously been argued that EF is a higher order construct that integrate these basic skills such as lower levels of attention. Within this framework, arousal and attention are acknowledged as foundation level skills, however EF is placed at this same level, suggesting attention and EFs are fundamental to the expression of skills above, thus denying that Luria’s (1980) unit one arousal and attentional skills actually underpin successful EF performance. The pre frontal lobe is a core area outlined in Luria’s third functional unit that synthesises other information that underpins executive reasoning and problem solving abilities, again contradicting the placement in Mapou’s framework. However, Barkley (1988) suggests complex attention and Executive Functions often overlap and are difficult separate, which might explain Mapou’s placement of attention and EF abilities together as foundational skills.

The first component of Mapou's hierarchical framework is global skills. This premise asserts that global measures such as general intellectual abilities are needed to underpin other functions. However, one would assume that you would need attention and arousal in addition to the Executive Function skills in order to complete many of the intellectual skills measured by standard IQ tests. With this view, Mapou asserts that the integrated skills at the top of the framework (where reasoning abilities are also subsumed) require the coordination of skills at the lower levels of this framework and argues that in-tact executive abilities are fundamental to learning and memory. He postulates that reasoning abilities and EF are difficult to separate, thus are placed on opposite ends of the framework. However, one could argue that the ability to reason is a core executive skill that could be placed at the top of the hierarchy within an EF framework. Thus, contention arises with the overall placement of EF in relation to other modalities, as many modality-specific skills are seen to overlap and the functionality of each is difficult to separate, further compounding the term. Perhaps a paradigm shift may be necessary in order to conceptualise the construct.

2.3.2 The schematic placement of Executive Functioning within a hierarchical modality framework

Mapou places the domain of EF lateral in foundation skills, however perhaps a hierarchical placement above in terms of complexity is better depending on the factors discussed previously. A schematic diagram of this paradigm shift is offered below (Figure 3) that implies a hierarchy of function, which is a concept well supported in the literature. Thus, not only does EF sit above other cognitive skills with respect to complexity, but also within the set of skills upon which an order of complexity may be imposed.

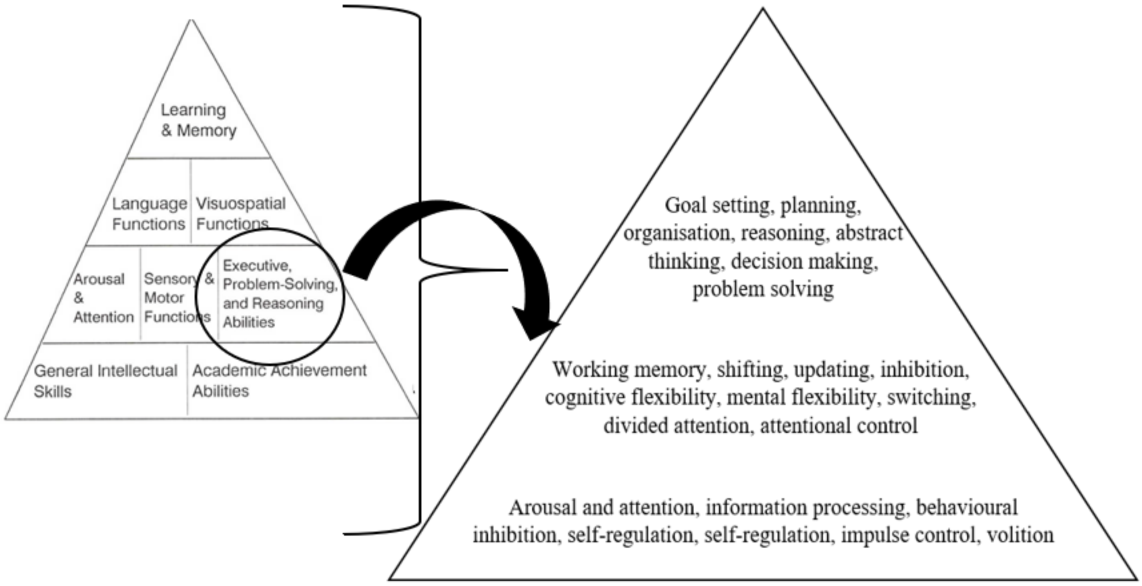


Figure 3. Schematic proposed diagram of EF conceptualisation.

Common to all definitions of Executive Functioning is the ability for one to carry out higher order skills for goal-directed behaviour (Alvarez & Emory, 2006), however identifying what these higher order skills are remains difficult to discern. If EF describes the ability to carry out higher order skills, then what are the foundation skills to EF, and are these necessary for the efficient functioning of EF? Beyond the consensus of EF being a higher order construct, there is lack of clarity regarding the constituents of EF and the differentiation between executive and non-executive skills. It is for this reason that discerning the interrelationships between those skills classified as executive further compounds the clear development of a model. This is because presentations of Executive Dysfunction (ED) can vary given that other cognitive abilities can be implicated, and executive control performance can be influenced from an array of non-executive skills (Lamar, Zonderman, & Resnick, 2002). Thus, as depicted in Figure 3 above, instead of placing foundational arousal and attention (essentially bottom up skills) in line with higher order EF skills such as reasoning and problem solving (essentially top down skills), EF may be better represented as a hierarchy of constructs that offers their own set of attendant skills that essentially underpin EF. For example, considering both a top down and bottom up approach may work in an integrative manner supported by distinct cortices to facilitate problem solving (or goal-directed behaviour), it may be sensible to consider the nuances of skills required for efficient Executive Functioning. For example, whilst arousal and attention, information processing, self-regulation, self-monitoring, and behavioural inhibition have been placed at the bottom, it is clear that these skills are drawn upon to effectively engage attentional control mechanisms such as cognitive flexibility, set shifting, updating and working memory, to facilitate planning for efficient goal-directed behaviour, reasoning and problem solving. Such a consideration could provide clarity with respect to how EF may then underpin other modalities such as language and visuospatial functions and learning and memory. Conceptualising the nebulous construct of Executive Functioning therefore remains challenging for a variety of reasons. Whilst some agreement has been reached regarding the overall applicability of the term, it is evident that little clarity exists beyond this point. Executive Functioning is a construct that refers to a *set* of skills, some of which are classified as an Executive Function (e.g., planning), and others which are not strictly classified under the umbrella of EF (e.g., working memory, attention, inhibition, and processing speed). Shallice (1979) proposed double dissociations as being indicators that cognitive systems are organised in a modular way where two tasks are handled by different modules. Interestingly, this is in line with the theoretical proposition of dual tasks, originally supported by Allport (1980). As discussed earlier in this chapter, dual tasks have provided much support for the operation of independent processing components either serially or in parallel, where a decrement in performance is observed if two tasks do not involve common processing components. Thus, cross-talk models of dual task interference suggest that interference may not depend on the type of task (operation) but instead proposes that it may depend on the content of the information that is being processed (Pashler, 1994a). Thus, there is evidence to suggest from a cognitive mapping perspective that there may be different systems where this may provide support for the operationalisation of independent modalities that work in an integrative manner for the smooth operation of EF overall, moving away from the original conceptualisation of EF as being unitary. For example, although controversial claims have been made, Fodor (1985) proposed that higher order thought processes such as reasoning and decision making do not operate in a modular way. Instead, weak modularity was proposed, which states:

> "although a particular network may be yoked to one particular cognitive function and receive one kind of input, some parts of the system may be involved in other networks as well (just as letters from assembled words in a Scrabble game can be used to build other words)" (Kosslyn & Koenig, 1992, as cited in Mapou & Spector, 1995, p. 245).

Thus, as implied in the definition of EF as a 'coordinator' that recruits a *set* of higher order skills, what if Fodor's premise applies to EF? Thus, the question remains whether EF is in actuality domain specific akin to language or visuospatial modules, and perhaps this is why to date, demarcation of the term remains elusive. However, a restriction of this premise is that Fodor terms each module as having no interface with or between other modules, and for something as complicated as EF that requires the integration of different modules (or constructs), a more sensible approach would be to consider EF with respect to the massive modularity hypothesis. Evolutionary psychologists propose that each module demonstrates functional specialisations that are highly interconnected and distributed across the brain (Frankenhuis & Ploeger, 2007). Given that the integrity of other modalities supported by a wider neural network is necessary for efficient Executive Functioning (Stuss & Alexander, 2000; Zelazo & Müller, 2002), it is clear that it relies upon other intact areas for input (Alvarez & Emory, 2006). This premise therefore seems reasonable to apply. Executive

Functioning must therefore be conceptualised in a way that is broad enough to capture the diverse anatomical structures that represents a variety of skills subsumed in EF (Alvarez & Emory, 2006).

2.4 Summary

As a construct EF must consider a) separate subdomains of EF skills, b) that subdomains be considered hierarchical and c) the directionality and relationship between each of the EF skills is evident. Empirical evidence has outlined that these features are paramount in which to classify EF and have been addressed throughout this review thus far. When conceptualising EF, perhaps "ownership" of skills does not demand exclusivity and delineating function based on modularity exclusivity is near impossible. Conceptualising the skills that comprise EF are near impossible given that a hierarchy may only be imposed and has yet to be validated empirically. Furthermore, dual tasks only infer that processing capacity is mediated by cognitive load serially or in parallel, which is also determined by the novelty and complexity of a task. Ultimately, the construct is defined by the skills.

Chapter 3 Operationalising the Skill set that Comprise Executive Functioning

Without argument, EF is referred to as a *set* of skills. Whilst the set of skills that comprise EF can help elucidate the overarching construct, it remains unclear what these skills are and how they are best assessed. Multiple skills have been proposed, however there is lack of agreement amongst theorists which undermines assessment techniques. Despite this lack of clarity, common to all is the promotion of the ability for one to function as an independent, self-serving adult (Lezak et al., 2004). Three commonly endorsed skills are labelled as shifting, inhibition, and updating (Miyake et al., 2000), however there is little agreement beyond this point. Many theories circulate within the literature and inconsistencies extend beyond what these skills are, but to the number of skills identified and the way in which to define them, often resulting in an overlap of purported skills. For example, in their comprehensive review, Jurado and Rosselli (2007) presented a range of putative skills proposed by numerous researchers, for example; volition, concurrent manipulation of information, effective performance, planning, cognitive flexibility, cue-directed behaviour, attentional control, information processing, goal setting, purposive action, impulse control, concept formation, abstract thinking, creativity, reasoning, problem solving, strategy generation, and purposeful coordinated organisation of behaviour. These are just a few of the skills purported by researchers, many often proposed more than once. However, another review of the literature demonstrates that there are over 68 different terms used to describe the skills comprising EF and over 98 executive tasks (Packwood et al., 2011). Clearly evident is that further research is warranted to offer a clear set of skills that comprise EF. However, the flaw in attempting to achieve this is that a plethora of skills have been proposed of which inconsistencies are inevitable, especially when they often overlap or contradict each other.

Perhaps a major contribution to the inconsistencies are the assessment paradigms from which these skills originated. Assessment paradigms are a distinct set of concepts to conceptualise a particular theory. Typically, EF paradigms stem from factor analysis studies that are used to reduce a large set of skills to a more meaningful and parsimonious number. However, assessment paradigms are also used to inform factor analytic research to identify the concepts to use in the first instance. Thus, a circular reasoning paradigm emerges, and inconsistencies surrounding the conceptualisation of EF is further compounded. Beyond the lack of consensus regarding the number of skills, disagreement regarding the taxonomies of EF further extends to the way in which researcher ‘label’ their factors, often resulting in an overlap and redundancy of skills when they are essentially the same. For example, shifting, mental shifting, cognitive flexibility, mental flexibility, working memory, divided and controlled attention have all been postulated in the literature, however it is unclear what their critical role is, or if they are the same or distinct. Is mental shifting the same as shifting of attention, and how or does this differ from cognitive flexibility? Thus, the way in which labels or factors are imposed by researchers is another contributing factor to the multitude of skills proposed above.

This chapter will discuss the assessment paradigms specific to EF, however results from factor analytic studies used to inform paradigms are considerably varied. Whilst these paradigms have been tested in a range of populations, a common restriction to conceptualising EF clearly is that both upward and downward extrapolation has occurred in relation to the age appropriateness of testing materials. Developmental trajectories differ between children and adults and EF skills come online at different ages. However, it is common practice to apply a test that may have been validated (within reason) in a particular age range, and then use this test in a different age range, where test requirements may not reflect appropriate developmental abilities. This in turn calls into question the validity of the psychometric properties of the tests of EF.

3.1 Developmental Issues

Developmental studies have shown that the brain continues to develop until early adulthood, and there are different trends in the onset, trajectory, and maturation of different EF skills over the childhood period (e.g., Anderson, 2002; Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Anderson et al., 2008; Baron, 2004; Best & Miller, 2010; De Luca & Leventer, 2008; Diamond, 2002; Diamond, 2013; Goldman-Rakic,1987; Grattan & Eslinger, 1991; Levin et al., 1991; Romine & Reynolds, 2005). Whilst a large body of literature has been devoted to understand the protracted growth and development of EF, developing tests that parallel developmental trajectories unfortunately remains challenging.

Current tests of EF for children are principally adult versions, which may or may not have been adapted to be more ‘child-friendly’. This is a pressing issue as adult measures are not designed to track development, and a child cannot be expected to perform the same as an adult on a skill that is not yet developed or matured. Clinicians however, tend to continue to use standard adult measures in the assessment of children, albeit with normative data derived for children. Conversely, using paediatric tests and extrapolating these upwards for adults may also present an issue for test selection and design since a complex task for a child may not remain complex for an adult if the skills being measured requires maturation of the frontal lobes. For example, the Day-Night task that is developed for children to measure inhibitory control relies on a child to say “day” or “night” when shown a picture of a sun or a moon. Clearly this task would be facile for an unimpaired adult, and not requiring inhibition or executive skills at all. Selecting EF tests that are developmentally appropriate therefore remains challenging and further compounds the development and use of appropriate tests. This is particularly evident when results from previous factor analytic studies and assessments paradigms illustrated below have used a mixture of ‘inappropriate’ tests (e.g., Tower of Hanoi (TOH); principally an adult version vs. Tower of London (TOL); principally a paediatric version) when drawing their conclusions about what a test is purporting to measure. Therefore, in both the paediatric and adult literature, there remains a failure to offer appropriate measures of EF.

3.2 Assessment Paradigms

Numerous paradigms exist that have given rise to many boutique tests that ultimately purport to measure the same thing but have failed to do so psychometrically in various populations (e.g., Tower paradigms). Whilst these measures all share the feature of a ‘problem-solving’ aspect, the requirements and level of complexity inherent in the tasks are considerably varied. For example, paradigms asses several cognitive abilities and multiple tests exist with varying clinical utility, strengths, and limitations. Whilst these paradigms provide support for an overarching domain, identifying impairment is difficult to discern as poor performance on a range of tests may include personality, past experiences, age, gender, and different strategies employed, that may not be restricted to pathology.

Fundamental to the principles of psychology are paradigms derived from theory and research methods, which represent distinct sets of concepts that allow the conceptualisation of a particular EF domain. Each paradigm has been strengthened and tested historically by utilising EF tests with varied populations and settings. Valid assessment of any EF skill therefore must respect the paradigm from which the test that assesses it originated from. For example, clinical assessment of planning is often achieved by adopting various tests from Tower paradigms, or visuo-spatial maze paradigms, and to a lesser extent, visual scanning paradigms. The purpose of the paradigms is to define parameters around specific skills or rules expected of the ‘patient’ to help elucidate the validity of the test and the performance outcome. Efficient assessment of EF will require the clinician to adopt numerous tests from numerous paradigms, to ensure a holistic assessment is achieved. This paper has already highlighted that cortical activation is widespread during many EF tasks, and thus selection of tests from only one or two paradigms should not be considered sufficient.

3.2.1 Maze paradigm

Maze (or hidden pathway paradigms) are commonly used as an assessment tool in cognitive neuroscience to measure EF and frontal lobe functioning (Milner, 1965; Walsh, 1994). This spatial learning task provides a useful basis for the understanding and the development of EF and spatial learning and memory (Bowden & Smith, 1994; Thomas, Reeve, Fredrickson, & Maruff, 2011) and has been used extensively in both research and clinical populations (e.g., Bowden & Smith, 1994; Crowe et al., 1999; Milner, 1965; Morrison & Gates, 1988; Tucker, Kinsella, Gawith, & Harrison, 1987; Walsh, 1994).

Although its original purpose was to detect frontal lobe impairment, it is clear that a task that assesses a spatial working memory element (particularly the navigation of a maze) may also recruit other cognitive skills related to Executive Functions, such as spatial working memory, error monitoring ability, visuomotor processing speed (Pietrzak et al., 2008), and planning (Bray & McDonald, 2010; Walsh & Darby, 1994).

More recently, defining the parameters around specific skills or rules expected of the 'patient' aided the validity of maze tests being used to measure EF. For example, within this paradigm, an individual is presented with a blank grid and required to find and learn a predetermined hidden pathway. There are certain rules that the individual must adhere to, and the hidden pathway is negotiated by one step (or tile) at a time by following these simple search rules. For example, rules permit movements of only one step (tile) at a time, and diagonal moves or retraction of the path are not allowed. A second trial then commences and the individual is required to find the exact same hidden pathway. However, disagreements exist amongst researchers as to what criteria are useful in identifying impairment on maze tasks, and therefore what maze paradigms purport to assess. For example, criteria for impairment on the Austin Maze test has been debated, as some suggest the inability to achieve 1, 2, or 3 errorless trials is indicative of impairment. Furthermore, achievement of various criteria may take a number of trials (e.g., 10-100), making it an impractical yardstick (Bowden et al., 1992). Beyond a stable error free performance indicating learning, the types of errors are also useful in identifying difficulties with memory, attention, impulse control or temporal integrations typically associated with frontal lesions (Fuster, 1989, as cited in Bowden & Smith, 1994). Thus, given the degrees of qualitative interpretation of error performance across a variety of studies, some have suggested that the maze may not be specific to executive disorders (Bray & McDonald, 2010).

Factor analytic studies using the maze paradigm have found differing results, and one of these is the Groton Maze Learning Task (GMLT). EFA found a two-factor solution in a clinical population, where total time and total maze efficiency (mean moves per second) loaded onto a Learning Efficiency factor, and legal, perseverative and rule break errors loaded onto an Error Monitoring factor (Pietrzak et al., 2008). Furthermore, Testa and colleagues (2012) found the Porteus Maze loaded onto a prospective WM factor with the TOL-R (time and correct), and Random Number Generation. Thus, whilst conjecture is evident, maze tasks require and therefore can be a valid measure of EF skills WM, feedback utility, monitoring and inhibition.

3.2.2 Card sorting paradigm

A widely accepted measure of EF (Greve, Stickle, Love, Bianchini, & Stanford, 2005), the Wisconsin Card Sorting Test (WCST) was originally a measure of shifting ability and abstract reasoning (Berg, 1948; Grant & Berg, 1948). The WCST has been used extensively in research and clinical practice as a measure of EF (Heaton, Chelune, Talley, Kay & Curtis, 1993) because of the array of skills necessary to complete the task. For example, the WCST measures a variety of complex skills such as abstract reasoning and mental shifting (Grant & Berg, 1948), and Heaton et al. (1993) suggest the WCST demonstrates a similarity to other measures of EF that require problem solving to achieve goals (Shallice, 1982). Furthermore, strategic planning, responding to environmental feedback, impulsive responding (Chelune & Baer, 1986; Welsh & Pennington, 1988), set shifting, and cognitive flexibility have also been proposed (Monchi, Petrides, Petre, Worsley, & Dagher, 2001).

This paradigm requires one to match or sort a deck of cards according to a sorting principle or rule, which is not revealed to participants. Four stimulus cards are placed in front of a subject with faces ranging in geometric shape, colour, and number, and feedback is given each time a card is placed down. Feedback is limited to telling the person right or wrong in their placement of the card, and no other information is revealed, especially when the sorting rule changes.

Traditionally, WCST performance has been attributed to the frontal lobes (Demakis, 2003; Milner, 1963; Monchi et al., 2001). However, caveats are advised when interpreting the WCST as a measure of 'frontal lobe functioning' because frontal patients have been found to outperform non-frontal patients (Anderson et al., 1991). Thus, contradictory literature emerges when attempting to identify the specific neuroanatomical substrates of the test (see Nyhus & Barceló, 2009 for extensive review of frontal and non-frontal performance on the WCST). Thus, the WCST requires multiple cognitive processes (Anderson et al., 1991) and whilst performance can be associated with the frontal lobes, it is clear that this is dependent on other cortical and subcortical areas (Mukhopadhyay et al., 2008).

Contradictory literature emerges when factor structures of the WCST are explored. Greve et al. (2005) reviewed 17 factor analytic studies on the WCST and found that most commonly reported is a three-factor solution typically assessing set shifting, problem solving/hypothesis testing, and response maintenance. However, a one factor solution has also been identified (Boone, Pontón, Gorsuch, González, & Miller, 1998; Bowden, et al., 1998; Swiercinsky & Hallenbeck, 1975). This may indicate that the WCST represents a single underlying construct, or aggregates of several processes that are inter-correlated (Greve, et al., 2005). For example, cognitive flexibility may be unique to the WCST where the subskills subsumed (e.g., set shifting, problem solving) are difficult to tease out, and might be why Boone at al. (1998) found a three-factor solution in 250 mixed participants that loaded onto a higher order frontal lobe factor. Furthermore, Testa et al. (2012) found that once the WCST was factor scored, a portion of the tests (WCST-FA; number of correct responses, the number of categories achieved, and the number of conceptual level responses likely representing reasoning and problem-solving), loaded onto the factor which they labelled Task Analysis. Secondly, the other portion of the factor scored WCST (WCST-FB; trials to complete first category, failure to maintain set and total correct, and reflected abilities required for set maintenance, which likely represented set maintenance/ staying on task), loaded onto a factor they labelled as Self-Monitoring and Set-Maintenance. Disparate findings are most likely attributed to the large scope of methodologies. For example, differing populations, sample size, test version (64, 128), and most importantly, the large set of variables that the WCST elicits where most are redundant to each other when all included under one factor. Thus, although identification of the factor structure of EF remains challenging, it might be because of the multitude of skills that are drawn upon for completion of the task, thus making the WCST a putative measure of EF.

3.2.3 Fluency paradigms

The overarching feature of fluency tasks is the generation of numerous responses within a specified time frame. The researcher or clinician is able to change the stimuli, such as words, figures, or categories (names, animals, places etc). The clinical utility of a fluency task is bound to the normative data available for the stimulus type.

Of all fluency measures, the most significant normative data is available for Verbal Fluency tasks in both healthy and clinical populations. The Verbal Fluency paradigm involves word generation tasks that are commonly used in cognitive neuroscience, and are part of many neuropsychological test batteries in both research and clinical practice to measure one’s spontaneous production of words within a set of prescribed rules. Alternate versions exist (e.g., Homophone Meaning Generation Test; HMGT, Excluded Letter Fluency task; ELF, Uses for Common Objects task; UCO, as reported in Strauss, Sherman, & Spreen, 2006), however phonemic letter categories (e.g., FAS) and semantic categories (e.g., animals) are the most commonly reported in the literature. Both phonemic and semantic fluency tap into frontal structures, however temporal structures have also been identified for semantic fluency (Henry & Crawford, 2004a). This paradigm requires participants to generate as many words as quickly as possible in one minute following certain rules, ultimately measuring a variety of skills (Strauss et al., 2006). For example, some skills that are thought to be examined include Executive Function (Fisk & Sharp, 2004), vocabulary, updating, lexical access speed, inhibition (Shao, Janse, Visser, & Meyer, 2014), and four main retrieval process such as: “(a) activation automatically spreading from the cue, (b) self-monitoring of output to prevent repetition and error, (c) suppression of previously retrieved responses, and (d) generation of cues to access new names” (Rosen & Engle, 1997, p. 224). However, other skills have also been proposed because there are other ways in which to evaluate performance on the test. For example, Troyer, Moscovitch, & Winocur, (1997) evaluated the strategies one may use to complete the task where store (clustering) and search (switching) processes are used. The authors proposed that clustering involving generating words within a sub category which is typically an automatic process, where Bertola, Lima, Romano-Silva, de Moraes, Diniz, and Malloy-Diniz, (2014), suggest this is a measure of semantic memory knowledge. Switching involves shifting from one category to another and is thought to assess cognitive flexibility (Troyer et al., 1997) as this is more effortful, where Bertola et al. (2014) suggests this is a measure of Executive Functioning. Furthermore, one may struggle to produce words at the beginning of the trial, which is thought to measure problems with task initiation. On the other hand, one may produce the majority of words at the beginning of the trial with a steady decline in production which is thought to measure difficulties in task maintenance (Delis, Kaplan, & Kramer, 2001 as reported in Strauss et al., 2006).

Moreover, inconsistent factor loadings are evident. For example, Verbal Fluency has been considered within the context of mental speed, where Boone et al. (1998) found the FAS loaded onto a factor in which they labelled as Speeded Processing. Furthermore, research has also found that both letter and category fluency load onto a single factor together (Ardila, Rosselli, & Bateman, 1994; Hedden & Yoon, 2006; Unsworth, Spillers, & Brewer, 2011), however they have also loaded separately. For example, Unsworth et al. (2011) found semantic and letter fluency loaded onto the same factor, which was then related to two other main cognitive constructs of Working Memory Capacity (WMC) and Vocabulary. Similarly, Whiteside et al. (2016) also found both animals and phonemic fluency loaded onto the same factor together with the Boston Naming Task (BNT) and the Vocabulary subtest (WAIS) which they interestingly labelled as a Language factor as opposed to EF. The other factor was labelled as EF due to loadings from the WCST and TMT-B. However, using a more comprehensive approach where 8 different category versions and 8 different letter versions were used, Schmidt et al. (2017) found semantic and letter fluency loaded on different factors, yet still shared common variance. Thus, disparate results are partly due to an inconsistent number of fluency items used (e.g., 3 in some studies, vs 8 in others) and the number of other tests used in combination (e.g., Whiteside's study used other cognitive measures whereas Schmidt's study only used fluency measures). On a mathematical basis, this affects the way in which they 'hang' together. However, even when a more holistic approach to fractionate the underlying cognitive components of EF was adopted by Testa et al. (2012), their EFA on 19 different tests that found six different factors, found both letters and animals loaded together with the Six Elements test which they labelled as a Strategy Generation and Regulation factor.

3.2.4 Stroop paradigm

The Stroop paradigm is an influential measure of interference control, cognitive flexibility, inhibition, and selective attention originally developed by Stroop (1935). Different test versions are available that propose to measure an array of skills. For example, inhibition, cognitive control, selective attention, learning, reading, language, working memory, conceptual ability and speed of processing (Boone et al., 1998; Kane & Engle, 2003; Miyake et al., 2000; Stroop, 1935; Uttl & Graf, 1997). The Stroop paradigm has been investigated extensively in both research and clinical practice (e.g., MacLeod, 1991), and has been shown to be supported by the DLPFC and the anterior cingulate cortex, (Zoccatelli, Beltramello, Alessandrini, Pizzini, & Tassinari, 2010), prefrontal cortices, and some parietotemporal regions (Yun et al., 2011). The premise for this paradigm is the automaticity of reading, where an individual must suppress the urge to read the written colour word, and instead name the colour of the ink it is written in. Thus, one must inhibit their pre-potent response to read the word.

As a timed task, the Stroop test has been reported to load onto a speeded element factor in a number of studies. For example, Boone et al. (1998) found that it loaded together with other timed tasks (FAS, Digit Symbol) labelled Speeded Processing. Interestingly, they also found that their second factor labelled Cognitive Flexibility (where the WCST loaded) and the third factor labelled Basic/Divided Attention and Short Term Memory (where Digit Span, Digit Symbol, ROCFT percent retention and Auditory Consonant Trigrams loaded) all loaded together on a higher order frontal lobe factor with a mixed sample of clinical and control participants. Similarly, in a large sample (*N*=390) of 9-year-old children, time to name colours, words, and the interference score (colours-words) loaded onto a factor together which the authors labelled as Naming Speed (de Jong, & Das-Smaal, 1993). Moreover, de Frias, Dixon, and Strauss, (2006) found the interference score loaded onto a single ‘Executive Function’ factor together with the Hayling sentence completion test, Brixton test, and Colour trails test in a large sample (*N*=427) of older adults (55-85 years). However, Testa et al. (2012) found the Stroop, TMT B-A, and Contingency Naming Task loaded on a factor labelled Set-Shifting and Interference Management. This is interesting because Testa et al. (2012) used a comprehensive approach which included many of the tests used in Boone et al.’s (1998) study, yet they did not load in the same way. This could be because of the clinical population in Boone et al.'s (1998) study or as mentioned previously, the type and number of tests used affects the mathematical rotation, and therefore final loadings. Therefore, despite the Stroop being generally recognised as a measure of inhibition or cognitive flexibility, there remains uncertainty as to what specific skills the Stroop test is measuring.

3.2.5 Pegboard paradigm

The pegboard paradigm is most commonly tested using Tower tasks and used as an assessment in cognitive neuroscience to measure planning (Arnett, et al., 1997; Lezak et al., 2004; Shallice, 1982). More recently, research has also highlighted the role of inhibition (Goel & Grafman, 1995; Miyake et al., 2000), working memory (Goel, Pullara, & Grafman, 2001), information processing speed (Arnett et al., 1997; Bestawros, Langevin, Lalonde, & Botez-Marquard, 1999), and problem solving (Goel, et al., 2001). Its popularity has increased despite there being little agreement as to what this test is measuring in a variety of clinical and non-clinical populations. Variants of the Tower paradigms include Tower of London (TOL), Tower of Hanoi (TOH), Tower of Toronto, TOL-F (Freiburg version) and a Tower task as part of the D-KEFS (Delis-Kaplan Executive Function System), all ranging in different aspects of complexity, instructions, and different sized rings/discs. Poor performance on this task is typically associated with frontal impairments (Beauchamp, Dagher, Aston, & Doyon, 2003; Goel & Grafman, 1995; Liang, Shewokis, & Getchell, 2016).

The apparatus consists of a wooden peg board with three upright rods and four different sized discs or rings, where participants are asked to move the pile of discs one a time to an end configuration, following simple rules. The premise of this paradigm is for participants to look ahead to determine the order and number or moves required to rearrange to the end configuration (Lezak et al., 2004). Factor analytic studies provide conflicting results, particularly when different versions are purported to tap into different skills in a range of populations. For example, Miyake et al. (2000) found the TOH was related to inhibition, whereas Handley, Capon, Copp, & Harper (2002) found the TOH was related to spatial Working Memory Capacity in undergraduate students. An interesting study conducted by Welsh, Cicerello, Cuneo, & Brennan (1995) found that a small sample (*N*=37) of healthy adult participants understood that sub goal attainment necessarily required more time devotion to make less errors. Participants also noted that during the pauses, they mentally visualised the movements and whilst evaluating plans they kept in mind the consequences of individual moves. Despite this, however, the ability to recognise that planning is an important skill of the TOH was not significantly associated with quality of performance. This was a unique study as the researchers used questionnaire data in addition to performance on the Tower, where they were able to ask participants why they paused whilst completing the task. They concluded that WM, problem solving, and planning are required to mentally visualise the sequence, evaluate goals, and keep in mind the consequences of rules when participants paused.

Other versions of the test have demonstrated different findings. For example, the TOL-R (Revised version where items were reduced to increase its reliability) was found to load on a prospective WM factor in healthy adults (Testa et al., 2012) and a Planning factor in healthy adults for the TOL-F (Freiburg version) (Debelak, Egle, Köstering, & Kaller, 2016). Similarly, Levin et al. (1991) found the TOL percentage solved on the first trial loaded onto a Planning and Strategy factor in isolation in children. However, Welsh, Satterlee-Cartmell, and Stine (1999) explored the contribution of cognitive skills to both the TOL and TOH in a small sample (*N*=37) of healthy adults, and found that whilst the two versions correlated well with each other, they found WM and inhibition measures explained over half the variance in the TOL compared to the TOH, where none of the WM measures correlated with the TOH.

3.2.6 Barriers in obtaining a valid factor structure of EF

As briefly demonstrated above, problems exist with the numerous assessment paradigms used in the measurement of EF. Each test has been used in various populations, with differing outcomes and thus selection of tests from only one or two paradigms should not be considered sufficient when assessing EF, as if used in isolation cannot be used to conclude the status of a patients' Executive Functioning efficiency. It is therefore why clinicians typically adopt numerous tests from numerous paradigms, to ensure a holistic assessment is achieved. For example, poor performance on a Tower task may be attributed to poor working memory, or poor inhibition, initiation, or a variety of other less complex skills, even though a core feature of Tower paradigm tests is the assessment of planning and problem solving. Furthermore, span paradigms include tasks that require storage and processing akin to working memory abilities on a range of verbal and non-verbal tasks. Updating, recall, and sorting paradigms are used to assess one's attention and concentration, WM and to a lesser extent, visuo-spatial abilities. Thus, a suite of measures is needed in order to measure all elements of EF reliably. Although these measures all share the feature of a "problem-solving" aspect, the requirements and level of complexity inherent in the tasks are considerably varied and identification of a tests' purported measure remains challenging in various populations. It is for this reason that highlighting "ownership" of skills does not demand exclusivity and delineating function based on modularity exclusivity is near impossible. Whilst overall assessment paradigms have been postulated that guide clinical and research practice, there remain problems within these assessment paradigms used in the measurement of EF. The factor analyses used to reduce the large set of skills to inform these paradigms have produced different loadings and are limited in their sensitivity and specificity. Therefore, in both children and adults, and in clinical and healthy populations, there remains limited generalisability of a valid factor structure of EF, and a comprehensive battery of tests must be included as a gold standard single test of global EF is not feasible.

3.3 Fractionation of Executive Functioning

Researchers have attempted to overcome the disparate nature of EF by adopting factor analytic (Exploratory Factor Analysis; EFA and Confirmatory Factor Analysis; CFA) studies to identify the skills likely assessed by a given test. Studies, however, have incorporated a mixture of outcome measures and the proliferation of the taxonomy of EF creates further inconsistencies in obtaining a valid factor structure in both children and adults (Brocki & Bohlin, 2004; Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Levin et al., 1991; Bennett, Ong, & Ponsford, 2005; Fisk & Sharp, 2004; Miyake, et al., 2000; Miyake, Friedman, Rettinger, Shah, and Hegarty, 2001; Pineda & Merchan, 2003; Testa et al., 2012; Welsh, Pennington, & Groisser, 1991; Wiebe, Espy, & Charak, 2008). This perhaps contributes to a lack of consistency in research findings regarding the sub-skills assessed by different EF tasks. Confounding results are partly due to some studies employing factor analytic techniques to varying populations. For example, factor analytic studies have been adopted in healthy adults (Miyake et al., 2000; Pires et al., 2018; Robbins et al., 1998; Testa et al., 2012), clinical (Bennett et al., 2005), paediatric (Brocki & Bohlin, 2004; Lehto et al., 2003; Levin et al., 1991; Welsh et al., 1991), and mixed populations (Boone et al., 1998; Greve et al., 2005). Moreover, test selection has also contributed to inconsistent findings. Some include a limited number of tests to evaluate specific Executive Function (Boone et al., 1998; Greve et al., 2005; Miyake et al, 2000; Schmidt et al., 2017; Strauss, Thompson, Adams, Redline, & Burant, 2000), unequal test selection, specific or wide age ranges, or small populations, whereas some have included a large range of tests (Testa et al., 2012). The different statistical methodologies employed also create disparate findings (e.g., EFA, CFA, Structural Equation Modelling; SEM), and the different types of rotations used (e.g., orthogonal, oblique, varimax) all contribute to the number of factors generated, and the way in which variance is extracted. As a result, using various cognitive measures of EF, researchers have identified three (Miyake et al., 2000), four (Fisk & Sharp, 2004) five (Pineda & Merchan, 2003) and even up to 6 different factors (Testa et al., 2012) presumed to capture EF just in healthy adults alone. More recently, Pires et al. (2018) reviewed more up to date factor structures although predominantly in paediatric samples, and found anywhere between one and three factors, where common factors reported were either updating, inhibition, shifting, EF, or WM (see Pires et al., 2018 for a review).

3.3.1 Unity and Diversity of Executive Function: Miyake, Friedman, Emerson, Witzki, and Howerter (2000)

Miyake et al. (2000) conducted a CFA on 14 tasks to confirm the factor structure of only three often postulated EFs known as Shifting, Updating, and Inhibition in 137 college students. Their second and main aim was to examine how these three EFs contribute to task performance on a range of commonly used EF tasks in neuropsychological research (WCST, TOH, Random Number Generation (RNG), Operation Span task, and Dual task). They found the three EFs were correlated, however are clearly separable. They concluded that shifting contributes to performance on the WCST, inhibition plays a role in the TOH, both inhibition and updating contributed to task performance on the RNG, and lastly, updating contributed to operation span. Dual tasking was not found to load onto any factors.

Whilst the authors directed a shift in the fractionation of EF which resulted in many researchers following suit to find the same three-factor structure (see Pires et al., 2018 for a review), the use of five overarching EF tasks combined with a limited number of preconceived factors that these might tap into (e.g., no planning, organisation), when it is known that there are other EFs, has led to a lack of validity in their conclusions which the authors have acknowledged. For example, other studies and normative data have found the

Tower tests are known to be a measure of planning, problem solving, attention, working memory, self-regulation, self-monitoring, and processing speed (Arnett et al., 1997; Goel et al., 2001; Lezak et al., 2004; Strauss et al., 2006), and other factor analytic studies have found this to load onto a Planning factor (Welsh et al., 1995). Given that inconsistencies are evident regarding the Tower as a measure of planning or inhibition, the use of only one arguable planning test (TOH) potentially limited the scope of the analysis. That is because if a test is said to measure planning but other measures of planning are not included, then this will restrict the ability of the factor analysis to produce a “planning” factor, or any other factor. As there were arguably no other tests included to measure planning, the TOH could only be predicted by the “inhibition” factor. Restricting the number of tests in this way has influenced the outcome. Although they did not set out set out to look at planning as a factor, it is still important to note this limitation, as it is vital to consider this in further studies. It is because EFs are poorly defined and multifactorial that these results vary so much across different studies. Past factor analytic studies have found that there are at least six factors that are commonly found despite differing test batteries and sample size, and large age ranges.

Miyake et al. (2001) conducted another CFA on 167 young adults however this time for the purpose of examining the relationship among Working Memory, Executive Functioning, and spatial abilities. Interestingly, they referred to the TOH as a task of ‘planning and goal setting’ and not ‘inhibition’ in their earlier study, where they asked participants to use the goal recursion strategy, which is considered to maximize the involvement of Executive Functioning. Similarly, they used the RNG this time also as a measure of ‘Executive Functioning’ not ‘inhibition’ as in their previous study. These results illustrate that the same tests may load differently between studies depending on what measures are used and for what purpose, which in turn creates the proliferation of inconsistent purported measures of a given test and the factors that are thought to explain EF.

This could be because of the methodological approach to a study. For example, depending on what the aim of the study is (e.g., to investigate three common EFs, or the relationship to Working Memory and spatial abilities) and what tests are included, changes the way in which the tests vary together mathematically. However, it is most likely that circular reasoning implicates the researchers' decisions to use tests in a different way. This is not at the fault of the researcher, it is because of the inconsistencies of what a given test is purported to measure, which is why researchers examine these paradigms in the first instance. However, the attempts to bring clarity to EF only produces disparity, where perhaps researchers need to be more unified in their approach.

Recent conclusions from Miyake and Friedman's latest work (2012) have postulated these three commonly referred to EFs are useful in understanding individual differences in EF. Their new framework suggests EF may be explained by a common EF dimension (unity) that is shared by all constructs. They found inhibition leaves no unique variance thus falls under the common EF factor. As such, their newly developed model has gained insight into the various ways in which EFs can different amongst individuals. As one example, the authors have highlighted substantial genetic contributions to both the unity and diversity levels of the model, and that a period of developmental stability is evident between the ages of 17-23.

Recently, Karr et al. (2018) conducted a systematic review and reanalysis of latent variables analyses to identify the best fitting models, given that Miyake and colleagues (2000) three-factor solution, and their recent common EF factor (2012) has garnered much support (see Karr et al., 2018 for *extensive* review). The authors found that no 'best' model was identified, however found tentative support for a three-factor/nested model, implying a common EF dimension. However, of the studies reviewed by Karr and colleagues (2018) it was evident that the same limitations from Miyake and colleagues (2000) exist. In the adult samples (ages 18-59) reviewed in Karr et al.'s (2018) study, anywhere between two and five models, and nested models were evaluated, with the most commonly examined being the three-factor solution.

Thus, the same issues are faced when using a limited number of tests. Moreover, evaluating a limited number of constructs (i.e., three) leads to considerable limitations. Karr et al. (2018) highlighted the need for an inclusive list of EFs and the issues with test selection, and for future research to explore and evaluate new constructs, especially as other EFs have been proposed such as planning, problem solving, fluency and reasoning (Packwood et al., 2011).

3.3.2 Factor structure of Executive Function: Pineda and Merchan (2003)

Pineda and Merchan (2003) explored 100 healthy Colombian adults aged 16-21 and found a five-factor structure explaining an impressive 74.9% of the variance. Four commonly used measures of Executive Function were utilised consisting of the WCST, Stroop, Verbal Fluency and TMT. Factor 1 consisted of the WCST measures (total error, categories completed, non-perseverative errors, perseverative errors) and was considered to be Cognitive Organisation and Flexibility. Factor 2 comprised of errors on the Stroop test, and was thought to be a measure of Sustained Attention, whilst factor 3 was the time to complete the Stroop test, most likely representing Speed for Inhibitory Control. Factor 4 was the time to complete TMT-A and B representing Visual Motor Speed, and lastly, factor 5 was comprised of the Verbal Fluency tests with no interpretation given. Interestingly, the Stroop speed and accuracy measures loaded separately, which were considered inhibitory speed and sustained attention respectively. However, whilst the amount of variance explained is impressive, this could possibly be a result of the limited number of tests used (only four), meaning that each factor only comprised variables from one test. To illustrate, all Stroop errors loaded on factor 2, all Stroop time variables loaded on factor 3, both TMT A and B scores loaded on factor 4, both FAS measures loaded on factor 5, and lastly, perseverative, non-perseverative, total errors, and categories completed for the WCST loaded onto factor 1. However, the WCST measures are redundant because the total number of errors is the summation of both perseverative and non-perseverative errors, meaning that observations within this factor are not independent. Thus, whilst a large amount of variance could be explained by this factor analysis, it may not be completely representative of 'Executive Function' given the limited tests and variables used (e.g., different components of attention, no planning, goal setting, or other commonly used higher order tasks).

3.3.3 Age-Related Impairment in Executive Functioning: Updating, Inhibition, Shifting, and Access: Fisk and Sharp (2004)

Fisk and Sharp (2004) conducted a Principle Component Analysis (PCA) on a range of EF measures in a large sample (*N*=95) of adults aged 20-81, in an attempt to confirm Miyake et al.'s (2000) three-factor solution. They also included the role of information processing speed to investigate if this weakened the relationship between age-related declines in EF processes. Whilst they arguably confirmed Miyake et al.'s (2000) factor structure, they found an additional factor in which they labelled as 'Access'. Thus, a four-factor solution that accounted for 67.56% of the variance was reported. They found that Reading Span, Computation Span, Brooks Spatial Sequences, and Consonant Updating loaded onto factor one which they labelled as Updating. Furthermore, Random Generation (alphabetic) and Random Generation (repeat) loaded onto factor two which they labelled as Inhibition. They also found that Chicago Word Fluency test (letter S), Chicago Word Fluency Test (letter C), and Random Generation (redundancy) loaded onto factor three which they labelled as Access. Lastly, the Wisconsin Card Sort (perseverations and trails per category attempted) loaded onto a fourth factor which they labelled as Shifting. Interestingly, they found that age was associated with a decline in performance for all three factors that were proposed by

Miyake (shifting, inhibition and updating), but not for their new factor which they interpreted as 'efficiency of access to long-term memory'. Lastly, when accounting for Information Processing speed, they found that this removed most of the variance accounted for by age where it was no longer a significant predictor for Updating, Inhibition and Access. For shifting however, whilst Information Processing speed removed half the age-related variance, age remained significant.

3.3.4 Factor analysis on nineteen Executive Function tests: Testa, Bennett, & Ponsford (2012)

More recently, Testa and colleagues (2012) conducted a comprehensive study into EF using factor analysis that seems to be the most encompassing in addressing past methodological issues. The sample consisted of an impressive 200 adult participants and used nineteen EF tests. Results produced six independent factors which they termed Prospective WM, Set Shifting and Interference Management, Task Analysis, Response Inhibition, Strategy Generation and Regulation, and Self-Monitoring and Self-Maintenance. Since EFs are generally recognised as poorly defined, Testa applied a unique approach to try and overcome this in her test selection. She triangulated test selection by having each task rated by herself, together with an independent neuropsychologist, which was an improvement upon methods described in previous projects. Their four criteria included: (i) how commonly the tests were used, (ii) if the tests were favourably reviewed in the literature, (iii) the psychometric properties of the measures, and (iv) the skills that the tests were purported to assess. Their ratings ranged from -8 to 8, with the aim to include those tests that scored above four. However, their recruitment did not control for education, intelligence, age, or employment history. A further strength of this research was attention to detail in the selection of outcome measures so as to reduce interference from irrelevant skills in the factor analysis. For example, participants completing the maze tasks were not penalised for touching the lines, as this is typically a measure of motor impairment. Thus, Testa and colleagues' study is one of the most comprehensive with respect to a large sample size and large test selection.

3.3.5 Assessment of Executive Dysfunction: Bennett, Ong & Ponsford (2005)

Bennett et al. (2005) explored whether other measures besides the Behavioural Assessment of Dysexecutive Syndrome (BADS) would be sensitive to executive dysfunction in a clinical population. The study of 64 TBI patients ranging in age from 17-73 years old resulted in a four-factor solution explaining 49% of the variance. Factor 1 comprised of the Zoo Map test, Porteus Maze, TMT-B and TMT B/A and most likely represents sequencing and self-monitoring. Factor 2 comprised WCST percent correct and percent perseveration and Modified Six Elements that was most likely representative of cognitive flexibility and set shifting in response to novelty. Factor 3 included Modified Cognitive Estimates and Action Program, which likely assessed conceptual thinking, and factor 4 consisted of Key Search, Rule Shift, Tinker Toy test and COWAT, likely measuring difficulties with initiation and cessation and control of action. The authors interpreted their factors using previous work from Banich (1997) because of the difficulties associated with labelling factors given the large number of variables and small number of participants. It was interesting to note that this battery was developed to assess Executive Dysfunction (ED), with component subtests that are purported to measure planning and goal setting ability, however no planning or goal setting factors were found. Their results further revealed that few measures were significant enough in predicting scores on the DEX (Dysexecutive Questionnaire) as rated by occupational therapists and neuropsychologists, and suggest that a combination of other measures in conjunction with the BADS may be useful for detecting ED.

3.4 Summary

All of the barriers to the clear conceptualisation of EF have thus far hindered the development of a completely successful unified model of Executive Functioning. Models are paramount for the successful conceptualisation and operationalization of a construct. They describe not just what the processes are, but 'how' processes occur. Therefore, it is essential to develop a model that describes the mechanisms of 'how' these processes work in the mind, to bring clarity to the tests used to guide clinical assessments. Tests should therefore correlate more effectively to represent the skills likely being assessed.

However, EF tasks currently lack many of the features necessary for clinical utility and evidence based practice. Problems with reliability, validity, and the theoretical underpinnings of EF threatens the ability to conduct meaningful neuropsychological assessment. These issues are problematic and cascade a myriad of implications. The flaw of EF both as a theory and in its assessment is that clinicians tend to draw conclusions inappropriately due to the lack of clarity described above. Assessments of EF are affected in particular as the models of EF are essentially derived by circular reasoning. The conclusion informs the question, and the question is the answer. For example, "George has a problem with planning. I know this because I used a test of planning to diagnose his condition. How do I know it was a test of planning? Because people who have planning problems perform poorly on that task." Circular reasoning implicates the diagnosis as the tests that are used to inform this basis have few psychometric properties to support them.

Although based on sound mathematical principles, factor analytic studies also fall short for reasons outlined above, resulting in sometimes widely inconsistent structures across different studies. This limits the ability to find consistent patterns, and as a result, in both the paediatric and adult literature there remains uncertainty of what skills are being assessed by EF tests, which in turn calls into question the empirical soundness of models of EF. Therefore, cognitive neuropsychological assessment must follow an information processing approach that recognises that other areas of cognition and modules work in an integrative manner. Assessments should be systematic, iterative, and work on the basis of hypothesis testing, as performance on one test should, and will lead to, another question that should be explored (Kay and Franklin, 1995). Thus, it is imperative to refer to a cognitive neuropsychological model of assessment (Kay & Franklin, 1995). Thus, in order to provide clinicians with appropriate assessment tools, the neglected areas of research are fundamental to discuss, particularly the development and clarity of a valid model of Executive Functioning.

Chapter 4 Models of Executive Functioning

The branch of psychology that focuses on modelling the mechanisms of action of cortical activation is cognitive psychology, and in essence provides the “how”. For example, how components of a particular system interact and are presented in the mind (Kay & Franklin, 1995). Cognitive theory has significant application to clinical settings as an understanding of how information processing systems work, usually depicted via discrete components or modules in box and arrow diagrams (Kay & Franklin, 1995). Cognitive models predict patterns of impairment (Kay & Franklin, 1995), and allow clinicians to infer where and why a breakdown in the system has occurred, consider the implications of the break and/ or severity (e.g., if a break is lower, then chain of processing will most likely fail) and lastly, remediate issues or provide strategies to help with dysfunction (which can only be done effectively if we can accurately identify and locate impairment).

A valid cognitive model informs theory and therefore testing techniques. However, in order to create a valid model, psychometrically robust tests must be available to test the construct. As outlined in the above sections, tests may be developed prior to formulating a model, however the opposite can also be true, whereby a theoretical model may be proposed which would be used to inform the tests (Kay & Franklin, 1995). Therefore, any assessment tasks that do not have a sound theoretical underpinning or a valid cognitive model would fall short of explaining the human brain-behaviour relationship. Unfortunately, the cognitive construct EF lacks many of these necessary features. EF theory, assessment and diagnosis of dysfunction is plagued by the lack of formalism inherent in the models used to guide research and clinical practice (Solso et al., 2014).

This paper has highlighted thus far the numerous issues that plague the construct of EF. In particular, given that cognitive abilities integrate a range of other skills, components, or modules, efficient assessments of EF remain challenging (Kay & Franklin, 1995). Therefore, valid assessment of any EF skill must respect the paradigm from which the test that assesses it originated from. This chapter will argue that the most pervasive EF models that will be reviewed lack formalism and grounding in the literature. Whilst these models have made significant gains towards the identification of specific skills, components and modules that comprise EF, there remains a lack of cohesion with respect to a unified approach and agreement towards a holistic model. It is critical that research continues in the attempt to rectify this in order to rectify to enable clinicians to understand the inter-operationalisation of the construct, particularly the influence of both executive and non-executive skills.

4.1 Supervisory Attentional System (SAS; Norman & Shallice, 1986)

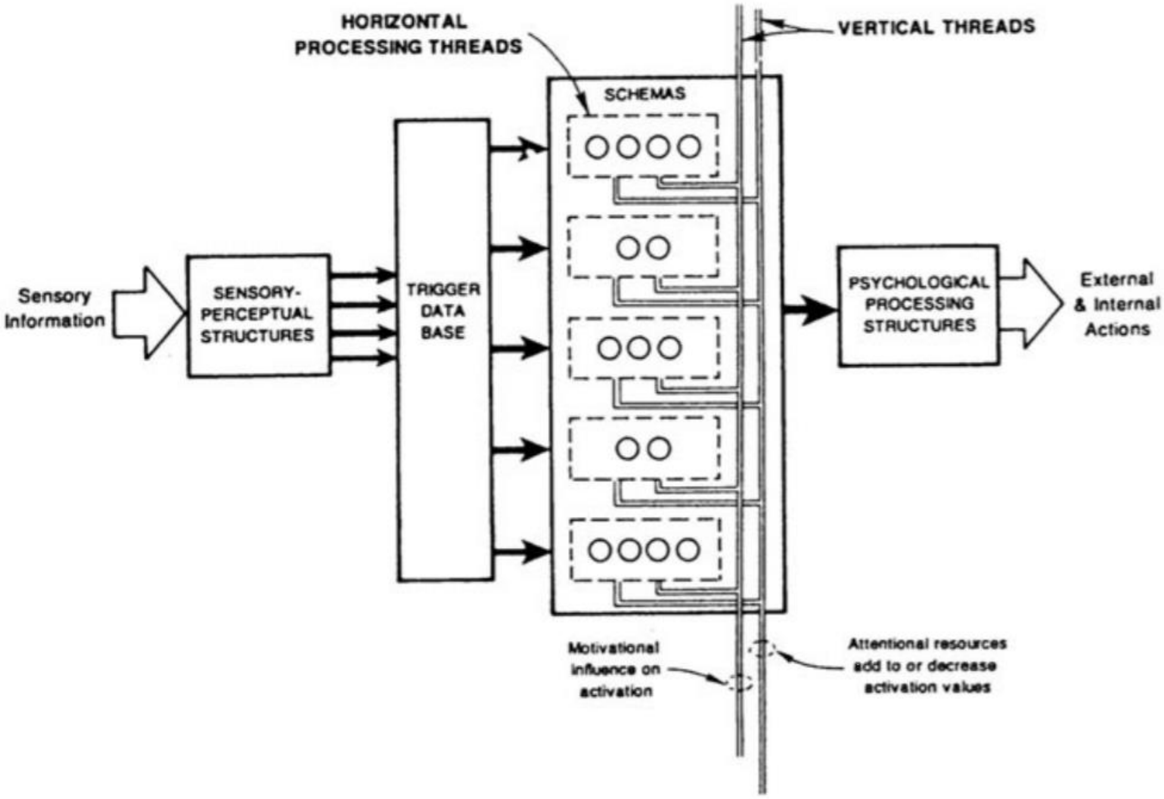


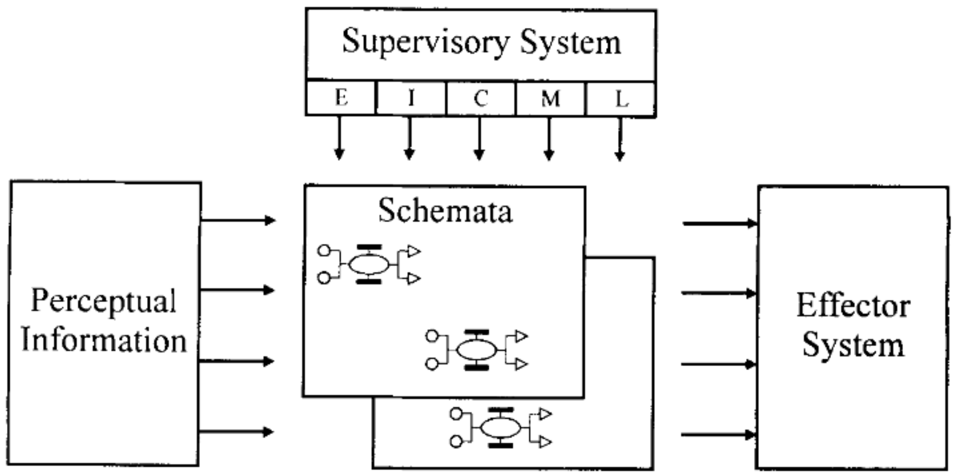
*Figure 4.*The overall Supervisory Attentional System. Reprinted from “Attention to Action: Willed and automatic Control of Behaviour,” by D.A Norman, & T. Shallice, in R. Davidson, R.G Schwartz & D. Shapiro (Ed.), *Consciousness and Self-Regulation* (4), p.7. 1986, New York, NY: Plenum Press. Copyright 1986 by Springer Science Business Media New York. Reprinted with permission.

The SAS is a model of Executive Function, originally developed by Norman and Shallice which highlights the role of attentional processes that modulate the selection of schemata, described as an implicit hierarchy. This model differentiated between those processes that are automatic (contention scheduling) and deliberate (SAS). Contention scheduling refers to automatic responses and its role is to inhibit and resolve conflicting schemata until an end state goal has been reached. During contention scheduling, the schemata present does not exist when a task is novel or complex (Norman & Shallice, 1986) since it has not been encountered before (at least in a literal sense). Thus, where a task is deemed novel and complex and additional attentional control is required (when planning or decision making is necessary), the SAS subsystem is activated. Thus, as implied in the definition of EF being activated when a task is complex, the argument that what is novel and complex for one may not be complex for another arises, thereby making it difficult to identify if a test is actually tapping into EF capabilities in a consistent way. Imaging supports the anterior regions, specifically the PFC, being active when the SAS is engaged (Norman & Shallice, 1986) supporting the contention of Luria (1966, as cited in Normal & Shallice, 1986) that 'deliberate attention' is linked to the PFC. Shallice and Burgess (1996) later expanded this model to purport that a globally integrated supervisory system can be broken down further to other subsystems involving the PFC. Specifically, three stages were proposed:

1. Strategy generation refers to new "temporary" schemata that is constructed for a novel situation. Problem solving is thought to play an important role, particularly with respect to the generation of new strategies. In order to generate new strategies, the involvement of problem orientation, goal setting, subsequent phases of attempting to solve a problem, as well as assessing to solution are required.
2. Working Memory is thought to be involved when the new temporary schemata is implemented
3. Schemata is monitored to either be rejected or modified

Whilst this model has been proposed as being a homunculus construct with a variety of subsystems all posited within the PFC, others have suggested that there is no Central Executive, and is best represented by the multifactorial nature of EF that requires a number of processes involving multiple brain regions. Specifically, Stuss, Shallice, Alexander, and Picton's (1995) *Supervisory System* was updated again to include anterior attentional functions, which exert high level Attentional Control consisting of five independent supervisory processes listed below, which can be viewed in Figure 5 below:

1. Energizing schemata refers to the activation and re-energisation if schemata become inactive (e.g., when requiring maintenance of attention)
2. Inhibition of schemata ensures inappropriate schemata are not activated
3. Adjusting of contention scheduling ensures that one given schema is not more active than others, and that successful inhibition of other unfavourable behaviour occurs
4. The monitoring system ensures there are minimal errors, the target schemata do not become inactive, and competing schemata do not influence behaviour
5. If-then analysis logic refers to the feedback of the previous stages to maintain and alter processes by reenergising, inhibiting schemata, or adjusting contention scheduling



E= Energizing schemata, I=Inhibition of schemata, C= Contention scheduling, M= Monitoring, L= “If-then logic”

Figure 5. Supervisory systems in human attention. Reprinted from "A multidisciplinary Approach to Anterior Attentional Functions," by D.T. Stuss, T. Shallice, M.P. Alexander, & T.W. Picton, 1995, *Annals of the New York Academy of Sciences, 769,* p. 193. Copyright 2006 by "John Wiley and Sons". Reprinted with permission.

Recent lines of work have supported the role of the SAS in a range of syndromes (Cooper & Shallice, 2000). Specifically, differential performance on tasks relating to inhibition or impulsive responding (conflicting schemata) between ADHD, Learning Disorders, and control groups demonstrated ADHD groups performed worse on inhibitory tasks (Bayliss & Roodenrys, 2000). This could perhaps provide support for the argument that foundation skills (contention scheduling) are critical for higher order skills. Numerous tests have been developed in accordance with this model, including the Six Elements Test, Hayling Sentence Completion Test, Brixton Spatial Anticipation Test, and sustained Attention to Response Task (Chan, Shum, Toulopoulou, & Chen, 2008). Furthermore, Wilson Alderman, Burgess, Emslie, and Evans (1996) developed a Behavioural Assessment of Dysexecutive Syndrome (BADS) to take into account dysfunctions that may arise in routine situations. Thus, whilst this model accounts for higher level Attentional Control by considering Executive Functioning, there remains issues regarding the applicability to assessment methodologies, given this is theoretical in nature (P. Anderson, 2008).

4.2 Lezak's Four Component Conceptualisation (1995; Lezak et al., 2004)

Lezak has attempted to classify the many ill-supported or agreed upon executive skills or 'capacities' in adults into four distinct categories. These are volition, planning, purposive behaviour and effective performance. She highlights that evaluation of the capacities that enter into the four EFs are paramount to the assessment of overall EF. She also highlights many skills are preconditions to some EFs which is an integral consideration given the theory behind EF. She has achieved clarity in organising these capacities into four distinct categories (e.g., formulating goals, planning, carrying out the plans to reach the goals, performing activities effectively). Lezak highlights that in order to carry out the activities it requires the capacities to "initiate, maintain, switch, and stop sequences of complex behaviour in an orderly and integrated manner" (Lezak, 1982, p. 290). One must also "conceive alternatives, weigh and make choices, and entertain both sequential and hierarchical ideas necessary for the development of a conceptual framework to the carrying out of a plan" (Lezak et al., 2004 p. 614). She also highlights that impulse control, memory and sustained attention are also necessary for the completion of a goals, where failure to complete a goal is because a failure of one or more of the capacities required to complete a goal are impaired (Lezak et al., 2004). Furthermore, the final category 'effective performance' requires the ability to "monitor, to self-correct, and to regulate the tempo, intensity, and other qualitative aspects of delivery" (Lezak, 1982, p. 293). Thus, it seems that the integrity of intact skills below in the chain of processing are necessary for efficient Executive Functioning (Alvarez & Emory, 2006), and these skills can be viewed in Table 1 below.

Table 1

Lezak's Four Component Conceptualisation of EF

Skill	Description	Skills assessed under this capacity	Outcome if impaired	Purported measurement
Volition	• Intentional behaviour • Determination of needs and wants • Required to formulate a goal	• Motivation and awareness	• Unable to think of anything to do • Complex task completion may be evident on demand with instruction	• Assessed separately by observations and questions
Planning	• Formulation of sequential and hierarchical ideas • Identification and organisation of steps to	• Sustained attention • Memory • Impulse control • Working memory provides insight	• Failure to achieve goals, possibly because of one or more of the abilities necessary to plan are impaired	• ROCFT[a] • Block Design • Picture Arrangement

	carry out intentions to complete a goal • Conceptualisation of changes, alternatives, and making choices	into planning problems due to the nature of maintenance of information to make a decision		• Maze and Tower paradigms
Purposive action	• Initiation, maintenance, switching and stopping in an integrative manner required for intention to plan • Contrasted to impulsive non-thought out behaviour • This 'programming' function is necessary for	• Self-regulation • Productivity • Mental flexibility • Shifting	• Poor regulation of performance may be evident if reduced productivity is demonstrated • Mental flexibility provides insight into self-regulation, where inflexible thought	• Slowed processing speed at the beginning or end of a task that is assessed qualitatively

	complex non-routine tasks, however not required for routine tasks (Shallice, 1982)		results in rigidity towards problem solving • Results in the inability to shift, in turn leading to perseverative behaviour and motor problems	
Effective performance	• Respond in a coherent and cohesive manner	• Monitoring performance • Self-correction and regulation of intensity and tempo of performance	• Difficulties responding in a coherent and cohesive manner	• The nature of the responses- particularly errors, distortions, and compensatory efforts

[a]=Rey Osterrieth Complex Figure Test

Whilst Lezak's conceptualisation has made gains toward identifying EF skills and proposing a sequence of stages and preconditions (in a latent hierarchy sense) of cognitive processing necessary for the overall performance of any given complex task, in addition to the specific tests used for assessments, there remains a failure to consider an attention component in the overall model. A further limitation of the model is the lack of bidirectional relationships between these categories when there is little disagreement within the literature that these skills relay critical information back and forwards between each other (Anderson, 2002; P. Anderson, 2008; Cepeda et al., 2013). Essentially, Lezak is missing 'how' these skills work in an integrative manner for effective performance. However, Lezak's conceptualisation is the most guided circulating within the literature to date, in that it is uncommon to read a paper that does not refer to Lezak when defining EF. This is likely because of the comprehensiveness of this conceptualisation that attempts to capture both cognitive and behavioural aspects of Executive Functioning, and the influence that one can hold on the other.

4.3 Behavioural Inhibition (Barkley, 1997)

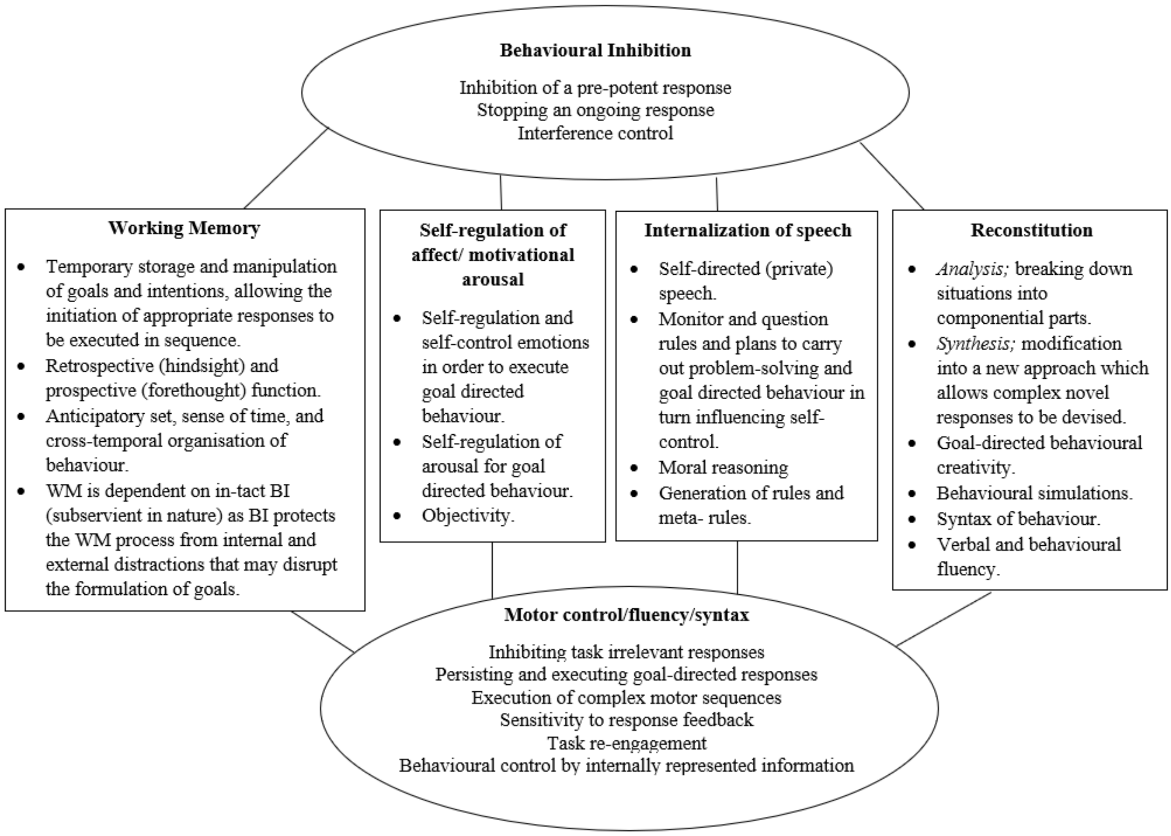


Figure 6. Barkley's (1997) neuropsychological model of behavioural inhibition. Adapted from "Behavioural inhibition, sustained attention, and Executive Functions: constructing a unifying theory of ADHD" by R.A Barkley, 1997, *Psychological Bulletin, 121*(1), p.73. Adapted with permission from personal communication with R.A Barkley, April 11, 2019.

Self-regulation is a response designed to alter the probability of an individual's subsequent response to an event (Barkley, 1997). Barkley asserts that self-regulation incorporates a large portion of the key components of EF as outlined by multiple definitions (e.g., goal-directed behaviours, planning, rules, impulse control). This hierarchical model places Behavioural Inhibition (BI) as a pre requisite, and proposes that BI provides a delay period which is a requirement before any executive process can occur. BI is made up of different interrelated processes which are (1) inhibition of a pre potent response, (2) stopping of a specific response pattern and (3) interference control which protects the delay period and executive processes. Thus, the model proposed by Barkley pays homage to vast literature surrounding self-regulation (e.g., Brownowski, Fuster) and the many other EFs that are largely dependent on this core aspect, and asserts that without the efficient function of BI, the remaining EFs suffer. This model has been based on ADHD studies which provides a link to the PFC and supports that inhibiting a pre potent response is a core manifestation of this disorder, which in turn affects sustained attention, which then reduces the ability to carry out efficient EF skills. In essence, one cannot sustain their attention if they fail to inhibit other distractions. Although Barkley suggests BI is distinct and hierarchically higher than systems controlling executive processes, others have argued inhibition is an executive process not a precondition (e.g., Miyake et al., 2000; P. Anderson, 2008). Barkley also acknowledges that a variety of validations are required for this model. For example, the strength of the relationship to other EFs, whether the specific subfunctions best represent each domain, whether these EFs can further be reduced, or if they are distinct, if they can be better represented as a general executive system, and lastly, if they are hierarchical in nature. Thus, whilst the current study agrees with a hierarchical representation of such subfunctions to comprise overall domains, particularly the pivotal role of BI and the level of detail regarding the evaluation of each of the other EFs, it does, however, agree that the strength and direction between all other EFs could both be expanded on to be included in a framework that includes higher order cognitive processing such as goal-directed behaviour, in addition to the influence that *all* domains have on each other.

4.4 Problem-Solving Framework (Zelazo, Carter, Reznick, & Frye, 1997)

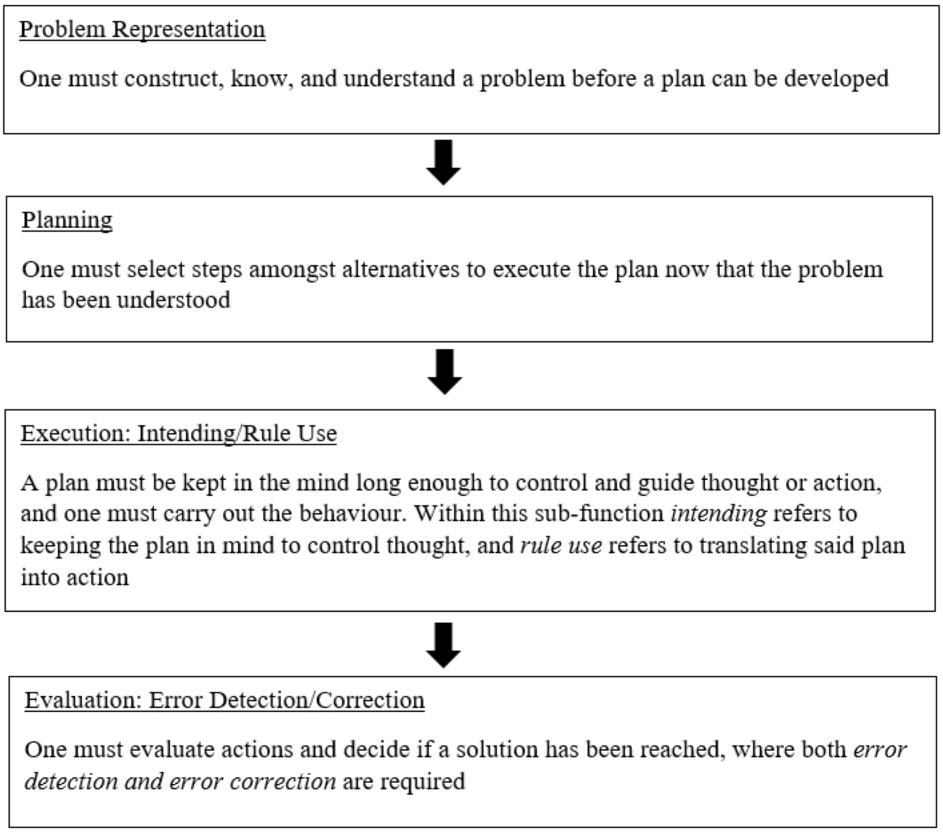


Figure 7. Zelazo and colleagues' (1997) problem-solving framework that identifies four temporarily distinct phases of Executive Function. Adapted from "Early development of Executive Function: A problem-solving framework," by P.D Zelazo, A. Carter., J.S Reznick., & D. Frye, 1997, *Review of General Psychology, 1*(2), p.200. Adapted with permission from personal communication with P.D. Zelazo, April 9, 2019.

Zelazo and colleagues proposed EF as a macro-construct, where the complexity levels inherent in self-regulatory processes required to problem solve are acknowledged. Furthermore, the proposed sub-functions are developed to highlight the way in which they work in an integrative hierarchical manner to achieve higher order function. The authors critique the way in which inhibition is represented by a 'basic function' that does not account for the complex 'metacognitive' processes involved in EF. Thus, a main strength of this model is the emphasis and incorporation of complex EFs, in particular, how they work in a hierarchical yet integrative manner to achieve a goal. Moreover, this model is able to determine where a break in the chain of overall processing skills is evident, and most importantly is the only model that highlights the outcome of Executive Functions; that is, this model is capable of demonstrating if the problem was solved or not, a common thing omitted in many other models. Although this model does operationalise Goal Setting, this model is less clearly defined. Thus, it is a further strength of Anderson's models that conceptualises Goal Setting as its own construct with its own set of attendant skills underpinnings its function that forms only part of the overall executive system.

4.5 *Anderson's Proposed Model of Executive Function (2002)*

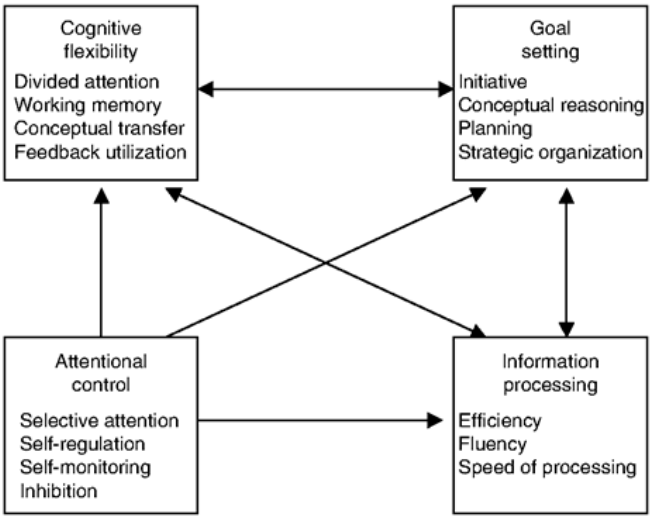


Figure 8. Anderson's (2002) proposed model of Executive Function. Reprinted from "Assessment and Development of Executive Function (EF) During Childhood" by P. Anderson, 2002, *Child Neuropsychology*, *8*(2), p.73. Reprinted with permission from personal communication with P. Anderson, April 9, 2019.

Anderson's (2001; 2002) developmental conceptual model has been derived from many factor analytic studies that aim to incorporate separate discrete subdomains that work in an integrative manner considering EF as an overall control system. Each sub domain comprises separate cognitive functions that share not only frontal systems but receive and incorporate stimuli from multiple regions in the brain.

Attentional Control; Selective Attention, Self-regulation, Self-monitoring, Inhibition

This subdomain involves selectively attending to specific stimuli where inhibition plays a role. Furthermore, this sub domain includes the ability to focus attention for an extended amount of time where the regulation and monitoring of actions are crucial to ensure plans are executed in the correct order and errors are identified early, so that goals can be achieved. Failure seems to include impulsiveness, impaired self-control and self-correction of errors, thereby leading to a lack of completion of a task.

Cognitive Flexibility; Divided Attention, Working Memory, Conceptual Transfer, Feedback Utilisation

Cognitive Flexibility includes the ability to learn from mistakes, shift, divide attention, develop different strategies, and, most importantly, process multiple sources of information simultaneously. Furthermore, working memory is also considered an important component within this domain. Impairments in this domain often reflect rigidity towards an approach to a task where one might struggle to adjust to new task demands and struggle to mentally hold and manipulate information, which in turn might lead to perseverative behaviour.

Goal Setting; Initiative, Conceptual Reasoning, Planning, Strategic Organisation

The description and skills purported under Goal Setting are in line with most definitions of Executive Functioning and is most closely aligned with Lezak's conceptualisation. The sequential order of actions that are strategic and organised are important components to consider, in that it alludes to a hierarchy of constituent skills, a commonly referred to yet an unclear proposition in the literature. This subdomain requires one to initiate an activity and devise a plan. To this end, it also encompasses the ability to plan, which involves anticipating future events and devising the steps necessary to achieve the goal. Organisation is key to successful performance, which includes organising the information in a logical a systematic fashion. Organisation is vital for how efficiently information is arranged, which makes it easier to retrieve information and plans. Implications of this domain include poor problem solving, poor planning, poor organisation, difficulties in developing strategies and poor conceptual reasoning.

Information Processing; Efficiency, Fluency, Speed of Processing

This is considered domain specific rather than globalised because of evidence from factor analytic studies suggesting response speed loads across different factors in children (Kelly 2000, as cited in Anderson et al., 2008; Welsh et al., 1991). Separate factor structures of processing speed are also well supported in adults (Chiaravalloti et al., 2003; Pires et al., 2008). This domain can be evaluated by the speed, quality and quantity of output, where this relies on intact frontal systems (Anderson, 2002). It is assumed that Cognitive Flexibility and Goal Setting cannot be processed without speed of output, efficiency and fluency, where this bidirectional relationship is vital because “performance on executive tasks can be significantly compromised in those individuals with slow Information Processing, however fluency and efficiency can also be enhanced when efficient organisational strategies are utilised” (P. Anderson, 2008, p. 17). Impairments may include reduce output, delayed responses and slow reaction times.

This theoretically driven model has been adopted and investigated empirically in numerous paediatric healthy and clinical populations (Bodimeade et al., 2013; Høie et al., 2006; Long et al., 2011), however, there remains a paucity of validation of the structure of this model in an adult population. Further, a strength of this model is that is has attempted to resolve the limitations of Lezak's conceptualisation in that it does have a clear attention component and acknowledges a bidirectional relationship between domains. However, a weakness of this model is that it fails to include an explicit hierarchy of skills (i.e., there is no start and finish in the overall efficient performance of a complex task).

Furthermore, Anderson's PhD thesis (2001) went beyond a conceptual framework and offered a set of assessments to examine each sub domain. Whilst a set of tests have been offered and is considered to be a strength of his discussion, there is confusion with respect to the skills that are purportedly being assessed, particularly when theory dictates that many of the skills overlap making it difficult to isolate skills to a particular sub domain. For example, aspects of Cognitive Flexibility and Working Memory are commonly tapped into when assessing Attentional Control because of the inherent nature of shifting attention (Anderson, 2001). Furthermore, many of the tests within Cognitive Flexibility are also useful for assessing components of the Attentional Control subdomain that tap into self-regulation and monitoring, and inhibition (Anderson, 2001). Lastly, Information Processing is a subdomain that assesses efficiency, fluency and speed of processing, however Anderson (2001) notes that few are specific to this domain, where reaction times serve as baseline data, and output production on switching, planning and reasoning tasks help evaluate Information Processing. Therefore, whilst a bidirectional relationship of the subdomains proposed in this model is considered a strength, the theoretical nature of EF prohibits single skills from being isolated (Miyake et al., 2000), which therefore hinders the clear conceptualisation of a model of EF.

Chapter 5 Rationale, Aims, Hypotheses

Of all domains of cognitive function, Executive Functioning has proven to be most vexing to researchers and clinicians alike. Whilst converging lines of scientific enquiry often narrow to a solution or bring clarity in research, it may be argued that EF demands a unified approach as it is itself a varied and complex skill-set. For example, psychological assessment has a long history of inferring normal functioning from abnormal pathology. Memory research has largely been grounded on the assessment of amnestic patients, including the infamous late H.M (Dossani, Missios, & Nanda, 2015). However, in EF research and assessment, abnormal pathology offers no significant “natural laboratory”, due to the widespread cortical activation associated with higher order function. The purpose of a model is to define and understand the parameters around cognitive performance and attempt to fractionate skills that contribute to performance, which would therefore work in parallel to begin to identify the function of certain cortices. Thus, to explore EF in a healthy population is not only beneficial, but essential.

To do so, effective assessment of EF requires that the executive tests themselves have good quality psychometric properties. This is necessary as it allows researchers and clinicians to understand and ensure that the skills and knowledge being assessed by the instruments are valid. However, one of the many major barriers to model development is that in the first instance, the tests that are purported to tap into EF skills are not psychometrically robust. One given EF test may measure a number of separate (yet perhaps related) EF skills. This in turn can affect reliability and validity because the task results can be viewed in different ways by different researchers. For example, one EF task can be classified as a measure of inhibition, planning, or even attention (e.g., Tower paradigms). Furthermore, some tests share skill similarities, thereby establishing convergent validity in psychometrics as demonstrated in correlation studies (Strauss et al., 2006). However, caution must be taken when interpreting these as ‘good quality’ tests because although statistical significance is achieved, correlations derived are often low. For example, two purported measures of ‘planning’, Key Search and Porteus Maze display significant yet low correlations (*r*=.28). Moreover, Word Fluency and the Ruff 2 & 7 selective attention test provide a correlation of (*r*=.17-.22). Lastly, the Hayling test has a correlation to the Brixton test *(r*= .14) when effects of age and IQ were accounted for (Strauss et al., 2006). These tests have significant but low correlations, indicating the confusion between tests. Perhaps the cross correlations between two tests are low because one test is not actually measuring what is it supposed to (Burgess & Shallice, 1996a) indicating low construct validity. To explain, two tests may purport to measure planning, however one test may not actually be a measure of planning, because it is unknown to an extent what the tests of EF are measuring, hence the low correlations between them. The overarching issue is that it is unclear which of the two tests actually best represents planning. It is not the notion of “planning” that is the problem, but rather that “planning” is likely a catch-all term for a set of complex skills, and each test may include or exclude elements in relation to each other. Essentially, the convolution of skills underpinning planning is likely demonstrating shared variance between skills, which is most likely why tests purporting to assess planning demonstrate low cross correlations. Thus, in order for researchers to successfully produce a coherent, effective, and psychometrically sound model and battery of EF respectively, researchers need to be unified in their approach to adopting numerous and varied EF measures.

This review has highlighted that factor analysis is a consistently utilised methodology. However, the differences in population, test selection, and statistical approaches employed, only serve to highlight that further consistency is necessary at the level of research design. Factor analysis dictates harsh demands in relation to sample size, power and the number of tests included for analysis. To that end, the validity of findings is questionable because theory is applied post hoc to mathematical findings, where relying on purely statistical support to fractionate an EF task that is well known to draw on an array of skills, only contributes to the proliferation of putative skills. Given EF tests can measure a number of separate yet related EF skills, defining uniquely separable skills is near impossible because tasks rarely tap into a single skill. Thus, task impurity obscures the clear conceptualisation of EF skills. Therefore, a comprehensive approach is paramount because without this consideration, justice is not given to the multifaceted nature of EF. It is therefore important to consider multiple measures of EF, particularly when model development is in question.

The aforementioned limitations restrict the clear conceptualisation of EF, which results in piecemeal studies. However, this is understandable because as already explained the harsh rules dictated by factor analysis restrict a comprehensive analysis, particularly within the cognitive neuropsychological literature. Because these are clinical tests, they must be administered face-to-face and often involve a variety of instructions under the guidance of the clinician or researcher. Therefore, the greater the number tests utilised, the longer the time to complete testing. Thus, a more appropriate approach, albeit ambitious, would be to use tests from varied assessment paradigms, which would inevitably dictate a larger sample size. This approach was adopted by Testa and colleagues (2012) in their study which used Exploratory Factor Analysis. However, it is clear that it is essential to be guided primarily by theory, with mathematics as a secondary concern.

Therefore, the crucial first step must be to clarify what the purported function of each test is and including more than one test for each function. Perhaps overlooked in previous research, or not appropriately emphasised is the value of assessing congeneric models in isolation prior to addressing a model in its entirety. Congeneric models allow theoretically supported, complex latent constructs to be analysed to examine the extent to which indicators represent a true generic score (Jöreskog, 1971). That is, they assume that tests measure the same latent trait- a variable that is not directly observed but are rather inferred (through congeneric modelling) from other variables that are observed and directly measured using neuropsychological tests. However, unlike other measurement models, congeneric models are not bound by such strong restrictions and therefore serve numerous purposes (Raykov, 1997a). For example, they allow different scales or measuring instruments to be used. Furthermore, a unique contribution of congeneric modelling is that it also addresses the importance of maximising individual reliabilities of composite scores that account for differences in weighting or contributions to the latent variable as well as accounting for error, thus providing an indicator of how reliably these scales represented the same underlying trait. In essence, congeneric modelling uses theory as a guide to test placement, and can be said to provide evidence of the construct validity of the individual items used to measure a particular latent trait (Jöreskog, 1971). The parallel argument therefore is that statistically, the use of Confirmatory Factor Analysis (CFA) via path analysis (SEM) to assess a full structural model of EF and the interrelationships between latent variables without assessing the validity of the tests it is founded on is flawed. Typically, an SEM approach is used to identify a series of tests, enter them into a statistical program, from which latent variables are identified and analysed for their relationships amongst each other. This approach often constrains the number of measures that can be included because this number is influenced by the number of participants. This approach also restricts the validity of the latent construct that is best represented by a selection of tests. This is because the emphasis is not on the placement of tests that represent an overarching latent construct (such as principle components analysis or factor analysis), but rather, the emphasis is on the relationship amongst these constructs, often neglecting the fundamentals of what is being used to measure the construct in the first instance. Therefore, it is essential to use theory as a guide for test placement.

5.1 Theoretical Rationale for the use of Anderson's Model

Anderson's (2002) paediatric model of Executive Function is selected as a framework to conceptualise EF in a healthy adult sample. This paediatric model addresses most of the limitations circulating the literature regarding a) how these skills contribute to the overall achievement of "independent purposive self-serving behaviour" b) proposes an inherent hierarchy of cognitive processes in the overall performance of EF, and c) the degree to which the directionality of these skills interrelate, thereby considering the theoretical nature of EF in a comprehensive way. It is therefore the strength of Anderson's model that adopts numerous tests from numerous paradigms to ensure a holistic assessment is achieved. This model presents an elegant solution that considers the integration and inter-operationalisation of an array of executive and non-executive skills within an overarching model of EF.

Although Miyake and colleagues' (2000) three-factor model has garnered much support, it is clear that it does not represent an exhaustive list of EFs (Karr et al., 2018). Thus, Karr et al.'s (2018) systematic review and re-analysis of latent variable studies highlight the need to consider both general and specific functions. Thus, although demanding exclusivity is near impossible, it is clear that a comprehensive inclusion of both EF and non-EF skills are beneficial when model development is in question. Therefore, the selection of Anderson's model was paramount.

For example, attention is a construct consistently omitted from adult models of EF, despite this being in most definitions. Evidence that attention must be considered when discussing EF comes from the fact that attentional difficulties manifest in many childhood disorders with an inherent executive dysfunction component (e.g., ADHD, autism, PDD), and shares functional attribution when an Executive Function is engaged (V. Anderson, 2008). The inclusion of an Attentional Control domain in Anderson's model is therefore more likely owing to the consideration of emerging skills in childhood, however, the consideration of attention in an adult model of EF is also paramount.

As a platform to begin this enquiry, the strength of Anderson's model is the consideration of selective attention, self-regulation, self-monitoring, and inhibition, and the way it shares similarities to Lezak, who proposed that one needs self-regulation and self-monitoring to regulate performance. A further strength of Anderson's model is the way it considers the influence of attention across different EF domains, and although not explicitly stated, the unidirectional arrows from the Attentional Control sub-domain suggests that this might be a fundamental mechanism to EF.

A review of literature suggests that many tasks may assess cognitive flexibility, and whilst some have found support for a factor in isolation, explanations regarding which skills may contribute to effective performance on cognitive flexibility tasks are lacking, and there are no models proposed to support this as a domain subsumed by a variety of skills. For example, previous research has proposed that a Cognitive Flexibility factor generally comprised the WCST variables (Boone et al., 1998; Bennett et al., 2005; Mirsky et al., 1991; Pineda & Merchan, 2003), however it does not necessarily validate the construct in its entirety. Anderson's model instead distinguishes Cognitive Flexibility as a sub-domain with its own attendant skills. This is a more satisfactory conceptualisation because it seems to be a 'catch-all' construct for a variety of skills underpinning its function. Where others have failed to conceptualise the distinct components subsumed by Cognitive Flexibility, the way in which Anderson includes divided attention, working memory, conceptual transfer, and feedback utilization under this sub-domain is a thorough conceptualisation and therefore considered a strength.

Another fundamental skill to all cognitive performance is the speed with which one can respond to task demands, and a key assertion throughout this paper is that speed and accuracy need to be considered when conceptualising EF. The speed-accuracy trade-off has been well documented where slowed processing speed is implicated when tasks increase in complexity, or vice versa, depending on goals, payoffs or motivation. Indeed, processing speed is a core component in psychological IQ testing and has been examined and identified as an influential component in not only EF, but also in a range of clinical populations. However, whilst many ‘information processing’ or ‘speed of processing’ factors have been proposed, rarely do we find a model with the inclusion of speed as a construct in the operationalisation of EF, or at the very least, both facets considered in task performance. As such, contention exists if speed of processing is domain specific or more globalised, which most likely comes down to the way in which processing speed has been measured previously. This may contribute to the lack of internal consistency that ultimately defines the construct as information processing, complex attention, cognitive speed, reaction time, or even psychomotor speed (Chiaravalloti et al., 2003). Thus, the unique aspect of Anderson’s model is that he proposes Information Processing as domain specific, bidirectional, and incorporating more than just the speed at which one completes a task, as opposed to an ‘ingredient’ function. Within this construct efficiency, fluency and overall speed of processing are proposed, where it is thought that Cognitive Flexibility and Goal setting cannot be assessed without Information Processing.

Lastly, EF comprises the ability problem solve, maintain and shift attention, inhibit pre potent responses, plan, implement various strategies, and utilise feedback, all of which are necessary to adhere to goal-directed behaviour (Royall et al., 2002). Goal-directed behaviour is therefore arguably the essence of Executive Functioning according to well accepted definitions. However, there remains a failure to offer a model of EF that captures Goal Setting as a construct, and a paucity of goal setting or planning factors that represents EF in its entirety (not just in isolation to a test). The models reviewed thus far consider different components of EF, however Goal Setting and its subsumed cognitive skills (e.g., planning, organisation) are typically the latent 'outcome', where models have never really explored the operationalisation of how one would actually set and achieve goals.

One model that attempted to operationalize planning or problem solving (i.e., setting goals) was Zelazo and colleagues' (1997) problem solving framework. Their model considers EF as a macro construct with hierarchical processes that are outlined to effectively plan to solve a problem. This model falls short however, because it lacks sufficient detail in relation to specific skills and their influence to effective problem solving. Lezak's conceptualisation of planning provides the most clarity as it includes consideration of a sequence of component skills necessary for its effective execution. Ultimately, the accumulation of skills at the lower level in the chain of processing skills are used to assess planning or Goal Setting. Furthermore, she considers that slowed processing speed at the beginning or end of a task may demonstrate reduced productivity and considers compensatory performance, where if one has inflexible thought then issues may arise when attempting solve problems. This may in turn reflect an inability to shift, which may lead to perseverative behaviour. It is therefore a strength of Anderson's model that pays homage to those that have attempted to classify Goal Setting, in that it instead demonstrates how this goal-directed behaviour is an executive as opposed to a binary outcome (based upon whether a person achieved the goal or not). Therefore, it is considered a strength that Anderson's model conceptualises Goal Setting to include the subskills of initiate, conceptual reasoning, planning and strategic organisation.

Together, Anderson's proposed sub domains are suggested to involve integrated cognitive processes sharing frontal systems that receive and process information from posterior, motor, and subcortical regions. Anderson (2001) further offers a battery of tests presumed to assess these cognitive processes, suggesting a microanalysis of assessment performance to enhance the "scoring systems" of each Executive Function test, to gain a comprehensive understanding of the multiple skills used for efficient Executive Functioning.

Therefore, as a platform to begin this inquiry, Anderson's paediatric model was used as a template for assessment for healthy adults. It is recognised however that children are not "mini adults", and adults are not "mini children," and that developmental trajectories differ. For example, most skills begin their protracted growth and come 'online' at different ages, and plateau around late adolescence to early adulthood.

For example, inhibition has been recognised as a non-uniform skill. Improvements have been noted between the ages of 5-8 and plateaus around late adolescence (Romine & Reynolds, 2005). Similarly, Stroop like inhibitory tasks have demonstrated gradual maturation until 21 years of age (Leon-Carrion, Garcia-Orza, & Perez-Santamaria, 2004). Working Memory has demonstrated linear trends in performance between the ages of 4-15 (Luciana, Conklin, Hooper, & Yarger, 2005), and planning as measured by typical Tower paradigms have demonstrated the more complex 4 or 5 moves reach adult level functioning around adolescence (Best, Miller, & Jones, 2009).

Thus, despite the inherent lack of appeal of extrapolating paediatric tests to adults key themes are still relevant for adults; for example, the way in which skills operate in a hierarchical manner, the similarities between top-down and bottom-up functioning, and the fact that multiple skills have been proposed to contribute towards efficient Executive Functioning. Thus, it is acknowledged that whilst the skills that comprise Anderson's model reflect paediatric level functioning, and that it is not sensible to replicate test utilisation exactly, every attempt was made to include measures that reflect the constructs using adult level normed tests.

5.2 Outline of Study Research Design

It is therefore the contention of this study that the existing models of EF cannot be clarified until validity is first established regarding its specific latent constructs, in turn creating validity within a given set of measures. Therefore, it is the impetus of the present study to confirm the validity of Anderson's paediatric model of Executive Function in a healthy adult population using a comprehensive approach, because it has been established to provide a template for assessment, yet has not been validated in the adult literature. This was achieved using a multi-step process. Firstly, in an attempt to avoid the methodological flaws of previous studies, a comprehensive approach to test selection was taken that mirrors (as far as is possible) the approach used by Testa et al. (2012). In order to achieve this, several tests from numerous paradigms were selected for inclusion in the overall study, with a triangulated approach to selecting good quality tests purported in the literature for healthy adults. Specifically, psychometric properties of approximately 50 tests were meticulously scrutinised in order to assess them against specific inclusionary criteria, and the lost was then reduced according to these criteria. A detailed correlational phase (variable selection and reduction phase) was also undertaken to further reduce the number of tests used.

The next step was to confirm Anderson's purported latent constructs using congeneric modelling. Clarification of test relationships achieved overall by the data reduction phase facilitated the assessment of each of the latent constructs identified in Anderson's model (Attentional Control; AC, Cognitive Flexibility; CF, Information Processing; IP, and Goal Setting; GS). Firstly, if Anderson placed a specific test under a certain construct, the current study matched this. Other tests were then placed as close to the theoretical placement recognised by Anderson. The value of the current approach is that placement of tests was underpinned by theory at the outset instead of being applied post hoc to mathematical findings. Doing so will offer a parsimonious, psychometrically robust battery of tests to best represent the constructs at each step of the model building process that other models have failed to consider.

The third step was to then assess the predictive validity and interrelationships between verified EF constructs. This was achieved through a series of regression analyses on constructs that were computed to create composite scores based on their reliability and validity as found through congeneric modelling. Doing so allowed for the assessment of directionality outlined in Anderson's model, and provided clarity with respect to how different skills contribute to the overall achievement of "independent purpose self-serving behaviour," which skills in the hierarchical chain of cognitive processes contribute more towards the overall performance of EF, and the degree to which the directionality of these skills interrelate.

5.3 Research Questions

At the broadest level this thesis is concerned with exploring the overall nature of EF, with respect to what skills it comprises, how best to assess them, and the manner in which latent constructs relate to one another to explain goal-directed behaviour. This study will explore:

1. Does a paediatric model of EF hold in a healthy adult sample?
2. How much variance in higher order complex measures of EF are explained by measures of attention?

By addressing this question, this thesis will clarify the contention in the current body of literature regarding the boundaries between EF and attention as separable constructs. Further to this, a question raised by current literature is the consideration of performance outcomes on EF tasks. Whether a task outcome is measured on a time scale or error based performance elicits very different interpretations of "good" or "bad" performance. A stronger explanation would be offered by considering "efficiency", a metric including both aspects. Thus, a final research question is:

3. Should speed and accuracy be domain specific, or considered within each task that is based on complexity?

Both clinical and research practice offer numerous means of assessing speed, and or accuracy, and further the notion of speed of information processing. These related concepts are often considered in the literature as a more generalised skill, subsumed and relevant to general cognition. However, it is important to highlight that measures of information processing speed are best measured by tasks traditionally considered EF (e.g., Fluency measures). Thus, the relevance of speed of task performance, how it influences error likelihood, and hence the overall efficiency, or speed of processing is an important aspect of EF that needs further clarity.

5.4 Aims of Present Study

Specifically, the aims of the present study were to;

1. Confirm the validity of Anderson's paediatric model of EF in a healthy adult population.
    a. Run four separate congeneric models on each sub domain to isolate the skills represented by Anderson's model (Attentional Control, Cognitive Flexibility, Information Processing Goal Setting)
2. Assess the predictive validity of Anderson's four subdomains
    a. Run a series of standard regression analyses where;
        i. Attentional Control, Goal Setting, and Information Processing will significantly predict Cognitive Flexibility
        ii. Attentional Control, Cognitive Flexibility, and Information Processing will significantly predict Goal Setting
        iii. Attentional Control, Cognitive Flexibility, and Goal Setting will significantly predict Information Processing

b. Elucidate lower and higher order skills in order to identify a hierarchy and explore the interrelationships between them.

5.5 Hypotheses

The EF constructs of Anderson's model are mostly bidirectional with the exception of AC which unidirectional and therefore hypotheses will only be outlined as such. Specifically, it is hypothesised that;

1) All four constructs (AC, CF, IP, GS) purported by Anderson will be upheld mathematically.
2) Attentional Control will be the strongest predictor, explaining the greatest variance in all other latent constructs. Thus, attention will be a significant domain that warrants its theoretical consideration within a model of EF, and not separate to it.
3) Information Processing will be the second strongest predictor of other latent constructs, therefore demonstrating that IP is an influential component in a model of EF.

5.6 Contribution and Significance of Study

The significance of the present study is that it will contribute toward our understanding of the differentiation of skills classified as executive by taking an approach that is underpinned by theory and supported by mathematics, rather than the other way around. Doing so will assist future researchers to develop a psychometrically robust battery of EF that will promote scientific validation and replication of EF within an adult population. This will facilitate both researchers and clinicians in test interpretation and assist them to conduct meaningful assessment with greater ease. It will also have significant clinical implications because such a list of tests can then be used to test theoretical models, from which diagnosis and treatment of ED and other psychological disorders can be improved. Therefore, the validation of a theoretically derived and robust model of EF in an adult sample will promote scientific validation and replication with numerous populations, ultimately bringing clarity towards the clear conceptualisation of the nebulous construct that is Executive Functioning.

Chapter 6 Methodology

6.1 Participants

Overall, 133 adults (42 male and 91 females) aged between 18 and 50 (*M*=29.68, *SD*=7.46) were recruited for the current study. The percentage of participants in each age bracket was: 18–19 years (5.3%), 20–29 years (56.4%), 30–39 years (27%), and 40–50 years (11.3%). Twenty-nine had completed or undertaken some secondary level education, 56 had completed or commenced tertiary courses, 28 had completed or commenced postgraduate studies, and 18 had completed or commenced a certificate/trade/diploma or other TAFE courses. Two participants did not report their level of education. The following criteria was used for inclusion, and only one participant was excluded as they did not satisfy criterion D. Specifically, inclusion criteria stipulated that:

A) Participants must be between the ages of 18 and 50

B) No diagnosis of any neurological, developmental or psychological disorders (e.g., autism or ADHD) must be currently present

C) Participants must not have undergone cognitive assessment in the previous two years, and finally,

D) Absence of significant intellectual impairment.

Criterion D was satisfied if participants demonstrated an IQ above 70. Participants were also required to self-identify if they had any known diagnoses of neurological disorders, however no participants were excluded based on this criterion. Mean IQ as a function of group membership is presented in Table 2, in addition to other key demographic variables.

Table 2

Demographic Characteristics of the Sample (N=133)

	Male			Female			Overall		Overall range
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	
Age	30.10	5.94	19-47	29.48	8.08	18-50	29.68	7.46	18-50
IQ	114.33	11.19	84-138	108.84	10.98	85-135	110.57	11.30	84-138

As indicated in Table 2 above, although the average IQ of the sample was higher than 100, the average score still fell well within one standard deviation (15) from the expected mean. Despite some studies suggesting there are sex differences in EF performance, the dominant consensus in the literature supports the contention that sex is not a significant factor to influence performance on EF measures (Strauss et al,. 2006), and therefore the sample was assessed as a whole, without separating male and female performance. Furthermore, participants in the study represented a fairly homogenous group with respect to Socio Economic Status (SES) (*M*= 66.17 *SD*=17.75) (ANZSCO range 8.9-100 with higher scores indicating higher SES). Tests were administered in visual and verbal counterbalanced order between participants order to minimise order effects. The number of participants that were administered each of the four variations were thirty-three in versions B, C, and D, and thirty-four for version A. The detail regarding test abbreviations and the order of test administration, are presented in appendix A and B respectively.

6.2 Materials

Psychometric properties of approximately 50 tests were meticulously scrutinised in order to assess them against specific inclusionary criteria, and a table of excluded tests is summarised in appendix C. First and foremost, if Anderson used a test, it was selected for the study (if available) even if it did not satisfy some criteria. For example, the ROCFT failed statistical criteria, however, was a test stipulated by Anderson and therefore was retained. Specifically, it was ensured that chosen tests for the overall study met the following;

i. Normative data available for the age of participants to ensure the current sample was performing within expected ranges.

ii. Include attention tasks that are putatively mediated by the anterior network (as discussed in Chapter 2). Specifically, measures of attention that are typically subsumed by posterior cortices (i.e. vigilance and orientation) were not included in the test battery.

iii. Produced outcome measures of speed and or accuracy, or offered the ability to flexibly adapt traditional scoring based on research evidence (as discussed in Chapter 2).

Tests were scrutinised for inclusion in the study not only to meet the criteria above but also to ensure there was a robust battery of tests in order to meet the demands of SEM. Thus, reducing a large battery of tests statistically by means of correlations for inclusion for further analyses was a crucial first step. This was necessary because researchers have emphasised the importance of providing a microanalysis of nuances of skills (Anderson, 2001). As such, many tests are scored in multiple ways including traditional according to standard scoring protocols, non-traditional according to alternative approaches denoted by other researchers, and derived scores according to the current study which included both measures of speed and/or accuracy. Applying this systematic approach will establish which tests hold strong relationships to each other and offer a smaller set of tests to confirm

Anderson's model, with an understanding of the most informative for inclusions into subsequent analyses.

Specifically, it was ensured that;

iv. Tests were construct driven, meaning at least three tests reflected each domain, in order to provide minimum coverage of a domain (Bagozzi & Yi, 2012; Hair, Anderson, Tatham, & Black, 1998; Hair, Black, Babin, & Anderson, 2014; Kline, 2015). Thus, in the instances where some tests failed inclusion criteria statistically, they were retained to provide a minimum coverage of a theoretical domain. This only occurred for a total of three tests that failed to meet statistical inclusion criteria, due to a failure to correlate above r =.3; the ROCFT (explicitly stated by Anderson to fall under the GS construct), Zoo Map, and Key Search (neither test explicitly stated by Anderson, however they were retained to provide a minimum coverage of three tasks per theoretical domain).

In order to provide a microanalysis of skills, tests were scored in varying ways. Specifically, tests were either scored;

a) Traditionally, according to standard test guidelines
b) Non-traditionally, according to alternative scoring approaches denoted by other researchers in the literature
c) Derived by the researcher, with the aim to combine speed and accuracy to consider both facets and obtain an overall efficiency score

One of the challenges of working with a normal sample in the way that was dictated by the current study is that EF tasks are designed to detect dysfunction, rather than to identify individual differences in performance. Hence, many of the mean results for some tests lacked variability, particularly those using an accuracy score. As a result, the assumption of normality was violated, meaning that these tasks could not be considered for the study. Those tests that did lack variability and did not violate the assumption of were retained. This however impacted criterion c above and the original aim to combine both speed and accuracy into an overall efficiency score was difficult to achieve because of the limited variability within some variables. Therefore, tests were considered using speed and accuracy as separate measures, with the aim to explore which measure would be more valuable during the data reduction phase of the study. A summary of included tests, the variables elicited, scoring formulas, purported measure and performance interpretation where applicable is available in appendix C. Those tests that were available in the public domain have their reported psychometric properties taken from Strauss, Sherman and Spreen (2006). Selected licenced measures were also included, such as the Behavioural Assessment of the Dysexecutive Syndrome (BADS) and the Test of Everyday Attention (TEA) that have been used extensively in both clinical and research settings and demonstrate consistent discriminant and construct validity. Listed below are all the objective measures proposed for use in the current study, where a total of twenty-two tests were administered.

6.2.1 Wechsler Abbreviated Scale of Intelligence: Vocabulary and Block Design Subtests (WASI: 2-subtest, Wechsler, 1999).

In order to estimate an abbreviated IQ, the Vocabulary and Block Design subtests (FSIQ-2) were administered according to standardised administration procedures outlined in the manual. The WASI provides short reliable measures of intelligence linked to the Wechsler Adult Intelligence Scale (WAIS-III) and the Wechsler Intelligence Scale for Children (WISC-III). The WASI is intended for ages 6-89 and consists of four subtests (Vocabulary, Similarities, Matrix Reasoning and Block Design). WASI IQ scores have a mean of 100 and a standard deviation of 15. Reliability coefficients for Vocabulary range from .90-.98, and .92 for Block Design (Wechsler, 1999). FSIQ-2 demonstrates reliability coefficients from .93-.98, test re-test stability of .88, and has demonstrated excellent validity (Wechsler, 1999). This test was used as a screening tool and to describe the sample. Scoring was triangulated where discrepancies were noted ensuring integrity of results. Administration time was 15-20 minutes.

6.2.2 Attention tasks

Trail Making Test (TMT)

This timed test is a measure of mental flexibility, speed, and attention for ages 15-89 years, originally developed by Partington & Leiter (1949), and later adapted by Reitan (1955). The test consists of two parts, part A and part B. In part A, the participant is asked to draw a line from 1-25 numbered circles organised randomly on the page, in sequence. Part B consists of the participant connecting 25 circles that are numbered and lettered, alternating between the two (e.g., 1A, 2B, 3C). Part B is purported to measure divided attention. Administration time for the TMT is no longer than five minutes, however no time limits were imposed, and a practice exercise was given for both parts.

Traditional scoring methods were applied, where the score for each trial was recorded as completion time in seconds (Strauss et al., 2006). Each participant was notified immediately if any errors were made and advised to correct their response, thereby adding time to their overall score. To obtain a purer measure of complex divided attention for part B (due to the differences in cognitive demands between the two trails) a derived score was also implemented (B-A; Lamberty, Putnam, Chatel, Beliauskas, & Adams, 1994). An attempt was made expand examination of errors in the present study to include self-corrected and instructor corrected errors, because of the influence each may have on speed and accuracy for analysis. However, data yielded no variance for error scores and only the following variables were used 1) TMT-A 2) TMT-B 3) TMT B-A, where a lower score indicated better performance.

Inter-rater reliability is reported at .94 for Trails A, and .90 for Trails B. Test re-test reliability is generally adequate, despite differing populations and age groups, ranging from .46-.79 for part A, and .44-.89 for part B (Strauss et al., 2006). Validity is demonstrated between A and B (*r*=.31-.6) which suggest they are moderately correlated purporting to measure similar, yet distinct, functions (Strauss et al., 2006). TMT has been linked with other measures of attention such as visual-spatial and scanning abilities, speed, executive control, cognitive flexibility and set-switching (Strauss et al., 2006).

Test of d2

The test of d2 measures selective and sustained attention, as well as speed of visual scanning for ages 9-60, developed by Brickencamp (1981). A series 47 of symbols are presented in fourteen lines down a page, consisting of the letters of ‘d’ and ‘p’ with either one to four dashes (‘) (“) presented above or below the letter. Participants are limited to twenty seconds per line, and are asked to mark all the letters that have a ‘d’ with two dashes; either two above, two below, or one above, one below. This task takes 4 minutes and 40 seconds to complete, not including instructions and a practice test. Traditional administration and scoring were utilised consisting of hits (H; maximum 300) misses (M), and false alarms (FA; maximum 358). A non-traditional concentrate score (CONC) was also used which is purported to measure speed and selective scanning and represents speed and errors equally (Bates & Lemay, 2004). As such the following variables were used 1) d2 H 2) d2 M 3) d2 FA 4) d2 total errors 5) d2CONC. A higher score for d2 H and d2CONC indicates better performance, and a lower score for d2 M, d2 FA, and total errors indicates better performance. Internal consistency is reported as *r*=.61-.97 (Bates & Lemay, 2004), and test-retest reliability is *r*= .89-.92 over 5 hours, and *r*=.92 over 12 months (Brickencamp, 1981). This test also demonstrates construct and discriminant validity (Brickencamp, 1981).

Stroop test

This test is a measure of cognitive flexibility, inhibition/interference, and selective attention (Strauss et al., 2006) originally developed by Stroop (1935) for ages 18-94. The Victoria version (Regard, 1981) was used, and materials extracted from Strauss et al. (2006). This version is ideal because of the ease of detecting response inhibition issues due to the shorter duration time, thereby reducing the possibility of practice effects. Participants for the Victoria version are presented with three 21.5 X 14cm cards, containing six rows of four items, spaced 1cm apart. Firstly, part D (Dots) requires the participant to name the colours of the 24 dots the card (blue, green, red or yellow) as quickly as possible. Secondly, part W (Words) requires them to name the colours of the ink used to print three common words (when, hard, over), and disregard their verbal content. Lastly, for part C (Colours) participants are required to name the colour of the ink used to print colour names (e.g., the word “red” is written in blue ink). All stimuli are arranged in a pseudorandom order, with each colour appearing once in each row, but used six times. Time to complete this task is approximately 2 minutes.

Traditional scoring was utilised (time in seconds), and errors were recorded (self-corrected errors were scored as correct). Participants were notified if their responses were incorrect if not corrected spontaneously. Furthermore, an interference score was also obtained (C/D; Graf, Uttl, & Tuokko, 1995). However, no variance was identified for the error scores and therefore only the interference score was used for analyses where a lower score indicates better performance.

Test re-test reliability has been demonstrated for the Victoria version (.90, .83, .91 for each part respectively), however practice effects have been noted with a slight increase on performance of 2-5 seconds (Strauss et al., 2006). Validity coefficients suggest moderate correlations have been demonstrated among test trials, indicating they are assessing similar yet different abilities as well as other measures of attention, inhibition, working memory and speed of processing and conceptual ability, as well as the interference score relating to a semantic system such as planning, and inhibition (Strauss et al., 2006).

The following tests are part of the Test of Everyday Attention test battery (TEA, Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994).

All measures from the TEA have been used extensively in both clinical and research settings (Robertson et al., 1994). As such they demonstrate consistent discriminant and construct validity, in ages 18-80 and are based on Posner and Peterson's (1990) model of attention.

Map Search

Participants are required to search for and circle symbols on a map of Philadelphia for two minutes, assessing selective attention. After one minute the pen is changed to discriminate number of correct items found across the first and second minutes. The score used for the present study was the number (out of a possible 80) found in two minutes, with a higher score indicating better performance. Reliability coefficients range between .80-87. This test loads onto selective attention in factor analyses.

Elevator Counting

This test measures sustained attention. Participants are asked to identify which level they have arrived at in the scenario that the elevator is not working. They need to identify which level they are at by listening to series of tape-presented tones. This test takes 2-5 minutes to complete. Traditional scoring rules were utilised where total number of correct answers were counted. Higher score indicates better performance, with a maximum possible score of 7. Results from this test were not directly used in analyses. Instead it was used as a base rate measure for the more complex test Elevator Counting with Reversal, and no variance was identified in this variable. Test re-test reliability coefficient is .88 for stroke patients.

Elevator Counting with Reversal (ECR)

This task is deigned to assess attentional switching and auditory verbal working memory. This task is identical to visual elevator, however instead of the arrow showing the directionality, an auditory high pitched tone is used to portray up and a low pitched tone will be used to portray down. Traditional scoring was utilised where total number of correct scores were counted with a higher score indicating better performance, with a maximum possible score of 10. This task takes 2-5 minutes to complete and test re-test reliability ranges between .66-.68.

Visual Elevator (VE)

This task is designed to measure one's attentional switching, where participants are presented a series of pictures with elevator doors. Small arrows show the direction of the counting (up or down), and participants are required to count the levels. Sometimes a vertical arrow will appear, and they should count this as a floor. For example, 1, 2, up, 3, 4, down, 3. Traditional scoring formulas as per TEA manual were utilised where two separate scores are derived. The first score is the number of correctly counted strings where a higher score equals better performance, with a maximum possible score of 10. The second score is the time to complete divided the total correct items, with a lower score indicating better performance overall. This test takes approximately 2-5 minutes complete and test re-test coefficient for this test ranges between .70-.79 for the timing score, and .71-.76 for the raw accuracy score.

Telephone Search

This task assesses visual selective attention and speed where participants are asked to look in a telephone book (scenario) for certain telephone numbers. There are two symbols (star, square, circle or cross) next to each service (plumbers, restaurants, or hotels), and participants are asked to indicate and circle when two of the same symbols are next to each other as quickly and accurately as possible. They are then required to tick a box at the end indicating they have finished their search. Traditional scoring procedures were utilised where the total time to complete is divided by the total correctly circles items (maximum of 20 correct symbols), with a lower score indicating better performance. This measure was not used individually in analyses but rather forms the baseline formula for Telephone Search while Counting. Administration time is 2-3 minutes and test re-test reliability for the raw score ranges between .86-.90.

Telephone Search while Counting (TSC- dual task)

This is a measure of selective divided and sustained attention. To complete this task the participant is asked to complete another telephone search while also counting a series of starting tones on a tape recorder (dual-task). This test score is calculated in a similar fashion as the telephone search described above (time divided total correctly circled symbols), and the dual task decrement weighted for the accuracy of tone counting is calculated by subtracting the time per target score from the previous subtest (telephone search), from the current time per target score for this subtest (telephone search while counting) (variable name= TSC-dual task). This task takes 3-4 minutes to complete, with a lower score indicating better performance. Upon inspection of the data the ceiling effects of the dual task decrement limited the range of data and hence reduced variability. This resulted in numerous extreme outliers that required truncation, however after this procedure reduced variability remained, with no inter correlations. Therefore, in order to retain a divided attention component, the non-traditional scoring method Telephone Search while Counting (variable name=TSC-E) was used instead, which does not attempt to control for the decrement and individual variation in processing or psychomotor speed however still captures a divided attention element. Test re-test reliability for dual task decrement ranges between .59-.61.

6.2.3 Span tasks

Block span: forwards and backwards (BS-Fwd, BS-Bk)

This task is the visual analogue of digit span (see next below), used to assess spatial attention (Smyth & Scholey, 1994) for forwards, and visual-spatial working memory for backwards (Wechsler, 1997b). It was originally designed by Corsi (1972) and Kaplan, Fein, Morris, and Delis (1991). Nine blocks are presented in a pseudorandom order positioned on a board. An examiner demonstrates tapping blocks in a sequence to participants whose task it is to repeat this sequence in forwards or backwards order. The number of blocks tapped in each sequence increases by one from a span of two to (up to) nine blocks. Traditional scoring protocols were used where the score is calculated by totalling trials correct for both forwards and backwards, with a higher score both indicating better performance (Kessels, van Zandvoort, Postma, Kappelle, and de Haan et al., 2000), and administration time is approximately 5 minutes. Block span was administered on an iPad with all administration and scoring wee computerised based on the original description by Corsi (1972) and scoring procedures by Kessels et al. (2000) (coding and graphic by Darby and Darby available from http://www.brainmapping.org/WhiteAnt/).

Average generalisability reliability coefficients calculated with Fisher's z transformation was .74 for forwards spatial span total score, and .72 for backwards spatial span total score (Wechsler, 1997b). Validity has also been demonstrated through a range of correlations amongst tests (Wechsler, 1997b). Brunetti, Del Gatto, and Deolgu (2014) found their e-Corsi digitised version did not differ substantially from traditional versions of the task, providing support for the computerized administration variant used in the current study.

Digit Span: forwards and backwards (DS-Fwd, DS-Bk)

Digit span is a common component to IQ tests such as the Wechsler Adult Intelligence Scale (WASI; Wechsler, 1997a). Participants are asked to recall sets of numbers in either a forward or reverse sequence. By doing so participants are required to mentally recall and/or reconfigure the order in which they were given, thereby assessing efficiency of attention for forwards (e.g., freedom from distractibility) and mental tracking such as working memory capacity for backwards (Lezak et al., 2004). Both depend upon short term memory capacity (Shum, McFarland, & Bain, 1990 as cited in Lezak et al., 2004). Digit span was administered on an iPad with all administration and scoring therefore computerised (coding and graphic by Darby and Darby available from http://www.brainmapping.org/WhiteAnt/). The researcher read aloud seven different pairs of number sequences, one at a time, at a rate of 1 number per second. Participants were asked to recall the numbers back (in ether forwards or backwards order), and the researcher tapped each number in the order they were reproduced on the iPad. Traditional scoring procedures were applied, and administration time was approximately 5 minutes. Variables used included trials correct for both forwards and backwards, where a higher score both indicates better performance. Average generalisability reliability coefficient for total score computed with Fisher's z transformation demonstrated to be quite good r=.86, test re-test stability is r=.83, and has demonstrated content validity between the WASI-R ad WASI-III r=. 83 (Wechsler, 1997b).

6.2.4 Fluency tasks

Verbal Fluency Test (FAS, Animals)

This controlled oral word search task requires the spontaneous production of words using strict search conditions intended for ages 18-74 (Strauss et al., 2006). Verbal Fluency has been used widely in research particularly from early work from Thurstone (1938) and thought to be a measure of Executive Function (Fisk & Sharp, 2004). There are different conditions of the Verbal Fluency test including; Semantic Fluency (aka animal) where the participant is asked to name as many animals as possible within a one minute time limit, and Phonemic Fluency where the participant is asked to name as many words as they can think of beginning with a certain letter (F, A, or S) within a one minute time limit. Participants are instructed to not repeat any words, not to use words that are proper nouns, and not to use the same word with a different ending (e.g., bed and bedding). Phonemic (FAS) takes approximately three minutes to complete, while Semantic Fluency takes just one minute.

Traditional scoring was utilised, with total admissible words counted. In addition, non-traditional scoring according to strategy based on store and search processes (Troyer, et al., 1997), and task initiation and maintenance (Delis et al., 2001 as reported in Strauss et al., 2006) was also used. Scoring the strategy processes is based on clustering and switching as proposed by Troyer et al. (1997). Clustering (or cluster size) refers to successively generating words with the same two first letters, differed by a vowel, or were homonymous for phonemic fluency and is thought to be a measure of verbal memory and word storage (Troyer et al., 1997) (Participants were advised in the present study to move on if two or more animals belonging to the same species were produced (e.g., red snake, brown snake, yellow snake= all snakes). Switching on the other hand, is thought to be a measure of strategy, essentially measuring cognitive flexibility and shifting (Troyer et al., 1997), or Executive Functioning (Bertola et al., 2014). Switching for fluency tasks was classified as the number of transitions between clusters including single words. To assess task initiation and maintenance, a line was drawn at each 15 second interval to denote a break-down of word production, as there have been differences noted in temporal allotments in word production (Delis et al., 2001, as reported in Strauss et al., 2006). The first interval is known to be automatic, and the latter quartiles are thought to be measure controlled processing (Hurks et al., 2006). Similarly, one may produce the majority of words at the beginning of the trial with a steady decline in production, which is thought to measure difficulties in task maintenance (Delis et al., 2001 as reported in Strauss et al., 2006). Lastly, the present study also derived scores to include types of rule break errors for analysis (rule break or repeat words), however no variance was identified for the types of errors. Lastly, derived scores were calculated for the average of the last three trials (e.g., 15-30, 30-45, and 45-60 seconds respectively). Therefore, the following variables were used 1) FAS-15 (total at 15 seconds) 2) FAS-AV (average of last three intervals) 3) FAS total correct 4) FAS clusters size 5) FAS switches 6) Animals-15 7) Animals-AV of last three intervals 8) Animals total. A higher score for all measures indicates better performance, however cluster size and switches are purely descriptive where a higher score was usually associated with better performance.

Internal consistency for FAS is r=.83 and test re-rest for both phonemic and semantic in healthy adults were above r=.70 in a range of populations (Strauss et al., 2006). Inter-rater reliability for clustering and switching scores were noted as greater than .95 for semantic and phonemic categories (Troyer et al., 1997). Validity coefficients range from .85-.94 for other phonemic fluency tasks, and animals ranging from 66–.71 (Strauss et al., 2006). Correlations with measures of verbal IQ have been noted (.44-.87) and is thought that episodic verbal memory also plays a role as reported in Strauss et al. (2006). Support from factor analytic studies purports an Attentional Control/Working Memory factor in both children and adults (Strauss et al., 2006).

5-Point test

This test is a nonverbal measure of production of novel designs as quickly as possible whilst following certain rules, in a manner similar to word fluency tasks, in essence measuring non-verbal fluency to assess executive control (Strauss et al., 2006). Regard, Strauss, and Knapp (1982) amended the original figural fluency version to overcome scoring and testing limitations, and was later adapted by Lee, Loring, Newell, & McCloskey (1994) to allow for a 3 minute time limit for ages 11 to adulthood, to make it comparable to the time limits of the phonemic (FAS) tests. All materials were extracted from Strauss et al. (2006), where participants are presented with a sheet of paper consisting of 40 dot matrices identical to that of a 5-dot arrangement on a dice and asked to complete as many different figures as possible in 3 minutes following certain rules. Rules stipulate that; i) only straight lines are allowed ii) all lines must connect dots iii) no figures are to be repeated, and iv) only single lines are to be used, with only one warning given for each of these violations.

Traditional scoring was utilised whereby total number of unique designs, total number of repeated designs, and a percentage of perseverative errors were also calculated ([perseverative errors/total unique design] X 100), as more productive participants have a greater propensity to make errors (Strauss et al., 2006), thereby assessing perseverative behaviour (Lee, Strauss, Loring, McCloskey, & Haworth, 1997). In addition to this, a non-traditional scoring method was used to incorporate number of generated designs in one-minute intervals, and a derived scoring method that included the average of the last two minute intervals, total number of rule break errors, and self-corrected errors, thereby including measures of both speed and accuracy. However, no variance was identified for repeat designs, rule break and self-corrected errors, therefore only the following variables were used 1) 5-point total Unique Designs (UD) 2) 5-point-1st 3) 5-point-AV 4) 5-point % perseveration errors (PPE). For all variables apart from 5-point % perseveration a higher score indicates better performance, and a lower percentage for errors indicates better performance.

Test-retest reliability is .78 for unique designs, and .51 for percent perseverative errors, and internal consistency is .80 for unique designs and .48 for percent perseverate errors in healthy adults (Fernandez, Moroni, Carranza, Fabbro, & Lebowitz, 2009). Thus, percent perseverative errors should be interpreted with caution, although low reliability is not uncommon (Lezak, 1995), and a range of explanations have been proposed (see Fernandez et al., 2009). Validity is evident demonstrating moderate correlations (r=.4-.7) with measures of visual spatial and constructive abilities (Block design) and executive control (WCST) as reported in Strauss et al. (2006).

6.2.5 Planning tasks

Austin Maze (AM)

Originally designed as a measure of Executive Function (Walsh, 1978), the Austin Maze is used to assess complex spatial working memory, planning, feedback utilization, speed of thinking, and spatial learning (Bowden & Smith, 1994; Bray & McDonald, 2010; Crowe et al., 1999; Milner, 1965; Stolwyk, Lee, McKay, & Ponsford, 2013; Tucker et al., 1987). The Milner (1965) pathway was used with instructions according to Walsh (1991). Participants are asked to select a hidden pathway, one block at a time, on a 10x10 grid of tiles. Hidden beneath the tiles is a 28-step pathway that leads from the bottom left to the top right of the grid. Participants are asked to follow simple rules; i) only one step is to be taken at a time ii) no diagonal moves are allowed, and iii) if an incorrect tile is pressed, return to the last correct tile and try again. It is anticipated that participants learn the pathway firstly by trial and error, but then avoid the incorrect tiles eventually through learning. Participants will see a tick or a cross depending on if they have found the hidden pathway or not, in addition to the auditory aid. Errors comprise selecting the same tile, backward, exploratory, diagonal, and skipped moves, fail to return, perseveration and double tap. Approximate duration is ten minutes.

Number of trials was limited to 10 as this has demonstrated a good estimate of error scores to criterion and learning in a range of populations (Bowden & Smith, 1994; Bowden, 1988, as cited in Bowden et al., 1992). These parameters have demonstrated cumulative errors over 10 trials is an effective method of assessment (Bowden et al., 1992) considering subjects with average IQ can complete the test in 10 trials (Tucker et al., 1987). To evaluate maintenance of error free performance, two error free trials was selected as this is the most commonly reported method in the literature (see Bray & McDonald, 2010), is most sensitive (Crowe et al., 1999), and research has demonstrated a high correlation between errors to criterion and errors over 10 trials in both normal (r = .89) and clinical populations (r = .94; Bowden et al., 1992). As such, the following traditional scoring methods were used 1) AM total errors to criterion (2 error free trials) 2) AM total time to criterion (2 error free trials). Of note, how many trials to criterion (2 error free trials) was violated with respect to extreme outliers and therefore was unable to be used.

The test was originally developed on a push button box but has since been adapted to multimedia platforms (Bray & Mcdonald, 2010), and was therefore administered using an iPad (coding and graphic by Darby and Darby available from http://www.brainmapping.org/WhiteAnt/). Research has demonstrated the equivalency of cognitive constructs being measured by both conventional and computerised versions (Morrison & Gates, 1988; Stolwyk et al., 2013) in terms of distribution of performance on both measures (McKay, Lee, Stolwyk, & Ponsford, 2012). For example, high correlations have been demonstrated between conventional and computerised measures (r= 0.74- 0.82), and no significant within group differences between the old maze and a computerized maze (Morrison & Gates, 1988). Given that equivalency of cognitive constructs has been demonstrated, psychometric properties can be assumed and interpreted. As such, internal consistency coefficient is r=.89 to r-.94, (Bowden et al., 1992) and construct validity has been demonstrated in the original versions (Bowden et al., 1992; Bowden & Smith, 1994; Crowe et al., 1999). Furthermore, test re-test reliability has been demonstrated r= .56 for trials to criterion (2 errors free trials), and r=.79 for cumulative errors to criterion (Tucker et al., 1987).

Tower of Hanoi (TOH)

This test is a complex measure of planning, inhibition, processing speed, problem-solving, working memory, visuospatial memory and problem solving (e.g., Arnett et al., 1997; Goel & Grafman, 1995; Goel et al., 2001; Miyake et al., 2000; Shallice, 1982). Multiple test forms can be used (Lezak, et al., 2004), and as such, the present study utilised an array of techniques, with administration time being approximately 10 minutes. The apparatus consists of a wooden peg board with three upright rods and four different sized discs mirroring the techniques and scoring of Bishop, Aamodt-Leeper, Creswell, McGurk, and Skuse (2001) so that more complex problems can be administered. Participants begin with a 3-disc trial, then a larger disc is added to create a 4-disc trial. Participants are required to move the pile of discs in varying sizes from one peg to another, following simple rule instructions. These are; one disc to be moved at a time, a large disc may not be placed on top of a smaller disc, and a disc may not be held in the hand or placed on the table while the other disc is being moved. A move was still counted if the participant took the disc from one peg to another, and then changed their mind without letting go of the disc. Participants were only reminded of the first violation of any of the rules.

The configuration of discs was adapted from Bishop et al. (2001) and the Psychology Experiment Building language (PEBL) software programme where a total of 13 trials were used altogether. One of the easier trials of the 14 from Bishop was removed to reduce the unnecessary repetition of easy trials (to be equivalent with the same trial removed from the computer-based version of the TOH). Given the healthy sample used in the present study, this ensured the most parsimonious administration of the task. Instructions and start and end configurations were presented on an iPad, and participants would then begin to solve the problem on the physical apparatus wooden peg board in front of them. Participants were instructed to complete all 13 trials.

Non-traditional scoring procedures as denoted by Bishop et al. (2001) were used. Using this method the final score was the highest level of task successfully completed (in terms of number of moves), with an additional half point added if both tasks at this level of moves were successfully completed (e.g., passed both 3 moves, 1 4- move, and both 5 moves= 5.5 (higher score indicates better performance). However, given the lack of agreement with respect to the likely skills being drawn on for performance of the Tower tasks, a number of traditional scoring methods were also used eliciting variables including 1) TOH Bishop 2) TOH time 3) TOH moves 4) TOH residual (how many moves over the minimum number did it take to complete;, a derived score) 5) TOH correct (correct counted as completed in minimum number of moves). Number of errors on the test violated the assumption of normality due to lack of variance, and was not used.

Internal consistency has been demonstrated at .87, and Chronbach’s alpha at .90 in a sample of 61 healthy adults for 12 trials working up to 15-move solutions (Humes, Welsh, Retzlaff, & Cookson, 1997). Furthermore, test re-test reliability for young children was over a 25-minute period was r=.71 (Gnys & Willis, 1991 as cited in Humes et al., 1997).

Rey-Osterrieth Complex Figure Test (ROCFT)

This test aims to measure planning, visual-spatial constructional ability, visual memory, organisation and problem solving strategies for ages 6-39 (Anderson, Anderson, & Garth, 2001; Lezak et al., 2004; Strauss et al., 2006). This test was developed by Rey (1941), and later adapted by Osterrieth (1944). Participants are asked to copy an abstract drawing presented to them as closely as possible. They are allowed to directly copy the figure in front of them for a minimum of 2 ½ minutes, and a maximum of 5 minutes.

All materials were extracted from Strauss et al. (2006). Traditional scoring methods were used including accuracy as delineated by scoring guidelines originally described by Osterrieth (1944), and later updated by L.B Taylor (1991) to include a stricter scoring criteria for ease of test interpretation. Thus, an 18 element system was developed with scores falling between 0-2 for each section of the figure according to whether sections of the figure were correct and placed properly, correct and placed poorly, distorted or incomplete but recognisable and placed correctly, distorted or incomplete but recognisable and placed poorly, and lastly, absent or not recognisable. A maximum score of 36 is obtainable. Furthermore, a traditional qualitative approach was also adopted utilising the organisational strategy denoted by Anderson, Anderson and Garth (2001). This system rates the drawing on a seven-point scale in terms of level of organisation (7- excellent, 1- unrecognisable) in both children and adults (Lezak, Howieson, Bigler, & Tranel, 2012) by switching the colour of the pen every 30 seconds to determine this score. As such, the following variables were used 1) ROCFT copy score 2) ROCFT copy time, a non-traditional scoring method 3) ROCFT organisational 4) ROCFT derived score (time/accuracy), where a higher copy and organisational score indicates better performance, and a lower score would indicate a better efficiency performance.

Internal consistency has been demonstrated by split half and alpha coefficients reliabilities great than .6 for the copy trial (Strauss et al., 2006). Practice effects are known to affect the ROCFT, therefore test re-test reliabilities are quite low (Strauss et al., 2006). Inter-rater scoring according to Taylor (1991) is high (>.90 as reported in Strauss et al., 2006). The new organisation score demonstrates excellent inter-rater reliability (r=.85-.92), and temporal stability (*r*=.79-.94) (Anderson, Anderson, & Garth, 2001). Construct validity has been demonstrated by correlation factor analytic studies for abilities including, visual-spatial organisation, visual perception, motor functioning, working memory (Strauss et al., 2006). Convergent validity has also been demonstrated (.12-.35) with traditional measures of EF (Anderson, Anderson, & Garth, 2001).

The following tests are part of the Behavioural Assessment of Dysexecutive Syndrome (BADS) test battery (Wilson, Alderman, Burgess, Emslie, and Evans, 1996)

All measures from the BADS have been used extensively in both clinical and research settings (Wilson et al., 1996). As such they demonstrate consistent discriminant and construct validity.

Rule Shift

This test is designed to measure the ability to shift attention. To perform this task participants are asked to say ‘yes’ to a red card and ‘no’ to a black card for 21 playing cards that are displayed in a spiral-bound notepad, turned over one at a time. A second administration of the task requires participants to say ‘yes’ if the card is the same colour as the previous card just shown, or ‘no’ if it is not. This test takes 1-2 minutes to complete, with traditional scoring utilised includes time taken, and total errors recorded. No variance was identified for errors so this was excluded, and Rule Shift 2 time was used over Rule Shift 1 as this was found to be a purer measure of shifting attention. A lower score indicates better performance. Inter-rater reliability is .98 for time to complete rule 2, and 1.0 for errors in control groups, and test re-test reliability is *r*=.-.08 in normal samples.

Key Search

This task requires a participant to plan efficiently and effectively and monitor their performance when asked to apply a search strategy to find ‘lost’ keys. Participants are shown an A4 piece of paper with a 10cm x10cm square in the middle with a small black dot 50mm underneath. They are asked to imagine that the square is a field where they have lost their keys. They are asked to show how they would search the field by drawing their search route, starting at the black dot. Administration times is approximately 5 minutes, and traditional scoring was utilised according to a marking system as delineated in the BADS manual that includes both speed and accuracy scores, with a total possible maximum score of 16. However traditional scoring only uses overall accuracy per search criteria, and as such the non-traditional time to complete variable was used. In addition, a derived score was utilised that incorporates both facets. The following variables were used 1) Key Search time 2) Key Search accuracy 3) Key Search derived (time/raw score). A higher score for accuracy indicates better performance, a lower score for time indicates better performance, and a lower derived score indicates better efficiency of performance. Inter-rater reliability coefficient is .99 in control and patient groups and test re-test reliability is r=.71 in normal samples.

Zoo Map

This task is designed to assess planning skills, and participants are asked to show how they would visit a series of location on the map of a zoo, following strict rules. There are two trails, and participants are asked to visit 6 places out of the 12, and are identical maps, but the rules vary. The first map is a high demand test that rigorously assesses planning skills in advance, to minimise errors. The second map is a low demand test where the participant is required to simply follow a checklist of instructions, therefore only the high demand Zoo Map 1 was used. Administration time is approximately 5-6 minutes, and traditional scoring instructions were utilised as per manual, as well as derived scores according to the current study. A decision was made to include self-corrected errors as true errors during this task, particularly as omitting these self-corrected errors would take away from the true nature of the task aimed at assessing one's planning ability. The following variables were used 1) Zoo Map planning time, a non-traditional measure 2) Zoo Map total time (total time- planning time for independence of observation between planning time and total time), a derived score 3) Zoo Map raw score, a traditional score 4) Zoo Map inversed raw score (to obtain same directionality of what constitutes a better score as a negative score can be obtained in this task which would indicate poorer performance), a derived score 5) Zoo Map derived score (total time/ inversed raw score). A lower score for total errors, and a higher score for raw score (where a possible maximum score of 8 can be obtained; number of correctly visited places – errors) indicates better performance. Lastly, a lower derived score would indicate a better efficiency performance. Inter-rater reliability is at .9-1.0 in controls and test re-test reliability is *r*=.39 in normal samples.

6.2.6 Reasoning tasks

Picture Arrangement (PA)

This task is typically used to assess concept formation, nonverbal reasoning and sequential thinking (Lezak et al., 2004). In healthy non-clinical samples this test is purported to serve as a nonverbal counterpart of the Comprehension subtest of the WAIS (as reported in Lezak et al., 2004). Picture arrangement forms part of the WAIS-III (Wechsler, 1997a), and traditional administration and scoring procedures were utilised. A higher score indicates better performance, with a possible maximum score of 22. Participants are asked to arrange a series of mixed up pictures displayed on cards to follow a logical sequence/story. Time to complete is approximately 5-10 minutes, with average reliability coefficient, *r*=.74, test re-test stability for a 16-29 age group *r*=.67, and the 30-54 age group *r*=.73, and lastly, demonstrated criterion related validity between the WAIS-R and WAIS-III (*r*=.63) (Wechsler, 1997).

Wisconsin Card Sorting Test: 64-card (WCST)

The WCST full 128-card version is the most commonly used measure of EF (Heaton et al., 1993). It is a complex task that draws on multiple cognitive abilities such shifting ability and abstract reasoning (Berg, 1948; Grant & Berg, 1948), problem solving to achieve goals (Shallice, 1982), strategic planning, responding to environmental feedback, impulsive responding (Chelune & Baer, 1986; Welsh & Pennington, 1988), set shifting, and cognitive flexibility (Monchi et al., 2001).

The 64-card version was utilised in the current study (Axelrod, Henry, & Woodard, 1992) in order to avoid participant fatigue (given the large number of tests administered). In this test, a participant is presented with four stimulus cards, and a set of 64 cards become their deck. The participant is asked to match each card with the four stimulus cards in front of them according to the principle they devise, but they are not told what the principle is (Colour, Form or Number). They are given feedback each time about whether or not they are correct, but not told why. The stimulus principle changes during task administration, however they are not informed about the change except that a previously correct rule may become incorrect, and they must adapt their sorting accordingly. This task takes approximately 10-15 minutes to administer, and traditional scoring procedures were utilised for the present study.

Traditional scoring practices were used as per manual instructions however, examination of the WCST variables indicated that trials to first category and failure to maintain set violated assumptions, particularly with respect to extreme outliers and normality which evidently was a result of lack of variance within these scores, and therefore they were not used in analyses. Further, due to the shortened version and the aforementioned violations, percent of conceptual level responses could not be obtained, and therefore was not used. The following variables were used 1) WCST total correct 2) WCST total errors 3) WCST perseverative errors 4) WCST non perseverative errors 5) WCST perseverative responses 6)

WCST categories completed. A higher score for total correct and categories completed indicates better performance, and a lower score for all errors indicates better performance.

A review from Greve (2001) suggests there are similarities between the full and shortened version, and therefore research from the full version can be cautiously generalised to the 64 card version. The same age-related decrements have been noted, along with the identification of poor performance on most variables in a range of clinical populations (Axelrod, Jiron, & Henry, 1993; Paolo, Axelrod, & Troster, 1996). Caution is only advised when extrapolating results from the short version to full for clinical importance (Axelrod, Paolo, & Abraham, 1997). Furthermore, research has demonstrated that majority of psychometric properties for the 64 card version are in keeping with the WCST full card version (Heaton, et al., 1993; Kongs, Thompson, Iverson, & Heaton, 2000) therefore as most aspects of the test are comparable to the full version, then the reliability and validity can cautiously be assumed, and all psychometric properties below are drawn from the full version manual (Heaton et al., 1993). Excellent inter-rater reliability has been demonstrated (r=.73-97), as well as generalisability (r=.39-.72,) and standard error of measurement (7.94-11.91). The WCST has been extensively used in clinical populations and research, and therefore demonstrates excellent validity (Heaton et al., 1993).

6.3 Procedure

Ethics approval was granted from the Victoria University Human Research Ethics Committee (VUHREC). Participants were recruited using a snowball sampling technique at Victoria University using flyers (appendix D) on staff and student notice boards (physical and electronic), as well as social media platforms (e.g., Facebook). Information was also placed in the global, Footscray Park and St. Albans Campus daily email bulletins. Advertisements included the phone number and e-mail addresses of the investigators. Any participant that was interested in taking part in the study contacted the PhD student investigator from the contact details listed on flyers or emails. Student investigators briefly explained details of the study in plain language, and information packs that included a plain language statement and information to participants, (appendix E) (including risks and benefits) were sent out via post or email. If interested in participating after reading the material in the information pack, participants were asked to contact the researcher to arrange time for testing to take place, and to read sign the enclosed consent form (appendix F) and bring this to their first testing session, along with completed questionnaires. All participants were screened for study inclusion by the PhD candidate, and whilst data collection was conducted through a team of researchers, the PhD candidate collected 30% of the data.

Testing took place over two sessions lasting approximately 90 minutes each, and sessions were spaced a minimum of one hour apart to minimise possible fatigue effects, but a maximum of 7 days apart to minimise drop out. The venue was either in a quiet room at the Victoria University campus, or at the participants' home, whichever was most convenient for them. Completed questionnaires and signed consent forms were collected at the first testing session. Tests were administered according to administration instructions, in a counterbalanced order between participants to minimise any possible order effects. A 5-10 minute break was allocated 45 minutes into each session.

For some tests it was unclear from the instructions whether or not certain strategies could be utilised. Strategies are a common tactic used in order to demonstrate efficient Executive Functioning, however, certain strategies take away from the true nature of the measurement of the test. Therefore, the decision was made to allow the use of strategies only in some circumstances. This was more so targeted at a particular test within the TEA battery, Elevator Counting with Reversal, whereby participants were not permitted to count on their fingers. Similarly, participants were not allowed to count on their fingers whilst completing the digit span tasks. With respect to the Zoo Map test of the BADS, in the instances where participants wanted to make markings on the paper to plan their route, they were not instructed they could do so, but were told they could, however only if they asked.

Chapter 7 Statistical Design and Data Reduction

It is the contention of the study to isolate tests represented by theory in order to clarify the latent skills that underpin EF tests. The parallel argument therefore is that statistically, the use of Confirmatory Factor Analysis (CFA) via path analysis (SEM) to assess a full structural model of EF and the interrelationships between latent variables without assessing the validity of the tests it is founded on is flawed. Therefore, a three-step approach was utilised each of which was guided by the former outlined in Table 3 below. For all analyses SPSS was used, and AMOS and process were additional software programs implemented where necessary.

Table 3

Three-step Approach of the Statistical Design of the Study

Steps	Purpose/description	Statistical approach	Assumptions met	Outcome
Variable selection and reduction phase	Data reduction on 57 variables based on criteria; 1) Theoretical relevance to Anderson's model of Executive Function. Specifically, if Anderson included a test in his model then a variable was included, even if this meant that criterion two was not satisfied 2) Statistical relevance a. Inter correlations were *r*=>.3 b. If *r* was not >.3 and the test was required to be retained for	Pearson's correlation	*Univariate Outliers*, *Normality*, *Missing Value Analysis (MVA) on all but TOH*[a], *AM*[b], *WCST*[c]	Total of 57 yielded 23 for further analyses, 34 removed

	theoretical purposes, then they were at least *r*=.25 c. Where multiple outcome measures were obtained from a single test, those that demonstrated a greater number of correlations were retained			
Confirmation of latent constructs	Confirmation of the four latent constructs of Anderson's model (Attentional Control, Cognitive Flexibility, Information Processing, Goal Setting)	Congeneric modelling, using Maximum Likelihood as the Estimation procedure	*MVA*[d] *on TOH, AM, WCST* *Sample size, Model misspecification, Model size, Multivariate normality, Estimation procedure, Linearity, Homoscedasticity, Multivariate outliers,*	Reliability, validity and composite scores obtained for all 4 validated constructs

			Independence of observation, model fit	
Predictive validity and interrelationships between constructs	Investigation of predictive ability of the relationships between validated constructs as identified in the previous step	Standard multiple regression	*Normality*, *Linearity*, *Homoscedasticity*, *Outliers*, *Multicollinearity*	All constructs contributed uniquely to a model of EF, some more than others

[a=]Tower of Hanoi [b=]Austin Maze [c=]Wisconsin Card Sorting Test [d=]Missing Value Analysis

Considering the unique aspect of the study utilising a variety of scoring methods, in order to reduce the data for subsequent analyses it was essential to apply specific criteria. Data reduction criteria were dictated according to the assumptions of SEM, which suggest correlations must meet r=>.3 (Hair et al., 1998), and at least three reflective indicators per latent variable are included to confirm a latent construct (Bagozzi & Yi, 2012; Kline, 2015). Thus, key selected variables were included for further analysis if they satisfied the criteria below;

1) Theoretical relevance to Anderson's model of Executive Function. Specifically, if Anderson included a test purported in his model then a variable was included, even if this meant that criterion two was not satisfied
2) Statistical relevance. Specifically;
    a. Inter correlations were r=>.3
    b. If r was not >.3 and were required to be retained for theoretical purposes, then they were at least r=.25
    c. Where multiple outcome measures were obtained from a single test, those that demonstrated a greater number of correlations were retained

Criterion 2b was warranted since some of the correlations were not above r=.3 for tests required to confirm Anderson's model. Hair et al. (1998) state that given that SEM is theoretically based, this warranted their inclusion. Furthermore, criterion 2c was difficult to satisfy at times and if violated, tests were only retained if purported by Anderson to confirm a construct. Again, this was warranted according to SEM requirements that specify at least three reflective indicators are utilised (Hair et al., 1998). A reduced set of variables could then be utilised for further analysis.

7.1 Data Preparation for Variable Selection and Reduction Phase Using Correlations

Assumptions were checked in reference to multivariate analyses below in the next section. All assumptions and statistical treatment methods for all stages of the research design are summarised in appendix G.

7.1.1 Univariate outliers

Various methods of handling univariate outliers exist, and therefore these were treated according to general rules of thumb (Orr, Sackett, & Dubois, 1991; Pallant, 2007; Tabachnick & Fidell, 2007; Zygmont & Smith, 2014). Orr et al. (1991) reviewed various treatment methods and found that some suggest not to remove outliers if they are legitimate, whereas others suggest it is necessary if they are extreme and there is a valid reason to exclude them. However, others suggest removing outliers based on extremity levels is not enough, and instead, the appropriate treatment should be to truncate or replace with the mean (Tabachnick & Fidell, 2007), or remove to ensure there are honest parameters in a data set (Judd & McClelland, 1989). As some scores were limited in their range (i.e., floor and ceiling effects) and thus variability, one or two deviations from the top and bottom indices resulted in sensitivity to extreme outliers (e.g., ECR, TSC dual task decrement) and as such, only extreme outliers denoted by boxplots were treated according to the aforementioned guidelines. This left non-extreme univariate outliers in the data, which were numerous (and in fact too many to transform) and were left because they were legitimate values (Orr et al., 1991). Approximately 50 values were truncated in the entire data set. Once raw scores were derived, extreme outliers decreased or did not occur.

7.1.2 Univariate normality

After derived scores were treated for outliers, all variables fell within the acceptable range of 3 for skewness and 10 for kurtosis, and therefore the assumption of normality was met (Kline, 2015).

7.1.3 Missing data

Acceptable cut off: SEM assumes there is no missing data however there is little agreement how much missing data can be tolerated (Tabachnick & Fidell, 2007). Some suggest 5% (Schafer, 1999), others suggest 10%-15% may result in bias (Bennett, 2001; Enders, 2003), whereas others still suggest if there is 20% missing, then 50% more data is required (Wolf, Harrington, Clark, & Miller, 2013). As such, there are no 'best' guidelines delineated because they are determined by a range of contributing factors.

Treatment methods: Various treatment methods to handle missing data have been proposed again with no consensus. Maximum Likelihood Estimation (MLE) is a widely used method for estimating SEM, and must assume multivariate normality (Allison, 2012). This approach is also equipped to handle missing data. Therefore, under MLE given the percent of missing data was small (<5%) for most variables (see appendix H), the model based impute method Expectation Maximisation (EM) was used as this the most commonly reported method in the literature (Allison, 2012; Cheema, 2014; Dempster, Laird, & Rubin, 1977; Dong & Peng, 2013; Schafer & Graham, 2002). This technique uses an algorithm to obtain the Maximum Likelihood Estimates (MLE). Thus, a Missing Value Analysis (MVA) was run on all variables besides the TOH and AM, because they were above the maximum requirement of 5% (22.6%, 15.8% respectively) (Schafer, 1999). Quality control checking of the data revealed problems with the administration of TOH after the first 30 participants, and therefore numbers for this task were reduced. For the AM, similar problems were identified for 20 participants also resulting in a reduction in sample. On the remaining data, Little's MCAR (Little & Rubin, 1987) (Missing Completely At Random) test produced a desired non-significant value; Chi-square= 378.060, df= 2132, p=1.000, indicating that data were in fact Missing Completely at Random. However, there is no guarantee in any data that missing points occur completely at random, therefore a more cautious approach would be to assume

MAR (Missing At Random) which is a more flexible assumption (McDonald & Ho, 2002), thus termed ignorable missing data (Allison, 2003). Essentially, if data are MCAR, then they are also MAR (Allison, 2012). Univariate normality and outliers were checked again, and the assumptions were met. However, closer inspection revealed that the estimation procedure affected the accuracy of each of the WCST variables (arguably acceptable percent missing; 7.5%), where the residual values of total correct should equate to total errors, however these values did not add up when replaced and therefore the WCST variables were not submitted to MVA in this phase.

7.2 Results and Discussion for the Variable Selection and Reduction Phase

An issue faced both by clinicians and researchers is that we are mired to traditional scoring methods, which ultimately affects the ability to assess or understand the clear picture of the nuances of skills. Thus, many researchers have suggested a microanalysis of assessment performance to enhance the "scoring systems" of each Executive Function test in order to gain a comprehensive understanding of the multiple skills used for efficient Executive Functioning. Variable selection and reduction yielded 57 variables for consideration. A bivariate Pearson's correlation analysis was run on all traditional, non-traditional, and derived scores to reduce the data to meet the requirements of SEM, and more than half were removed. The governing criteria for retention of a variable was the number of intercorrelations evident, and to a lesser extent the intracorrelations Ultimately, those considered the 'best' were those that either demonstrated a higher number of correlations above r=.3, or were pertinent to the constructs postulated by Anderson. It was revealed that most variables correlated to a greater extent with variables from the same test-set (intracorrelations), in comparison with the number of correlations across different tests (intercorrelations). All correlations that were r= .3 and higher were significant, and in some instances those just below r= .3 were also significant. Figure 9 below displays the frequency distribution of all 57 variables, highlighting those retained for further analyses, and descriptive data can be viewed in appendix I, and correlation summary found in appendix J.

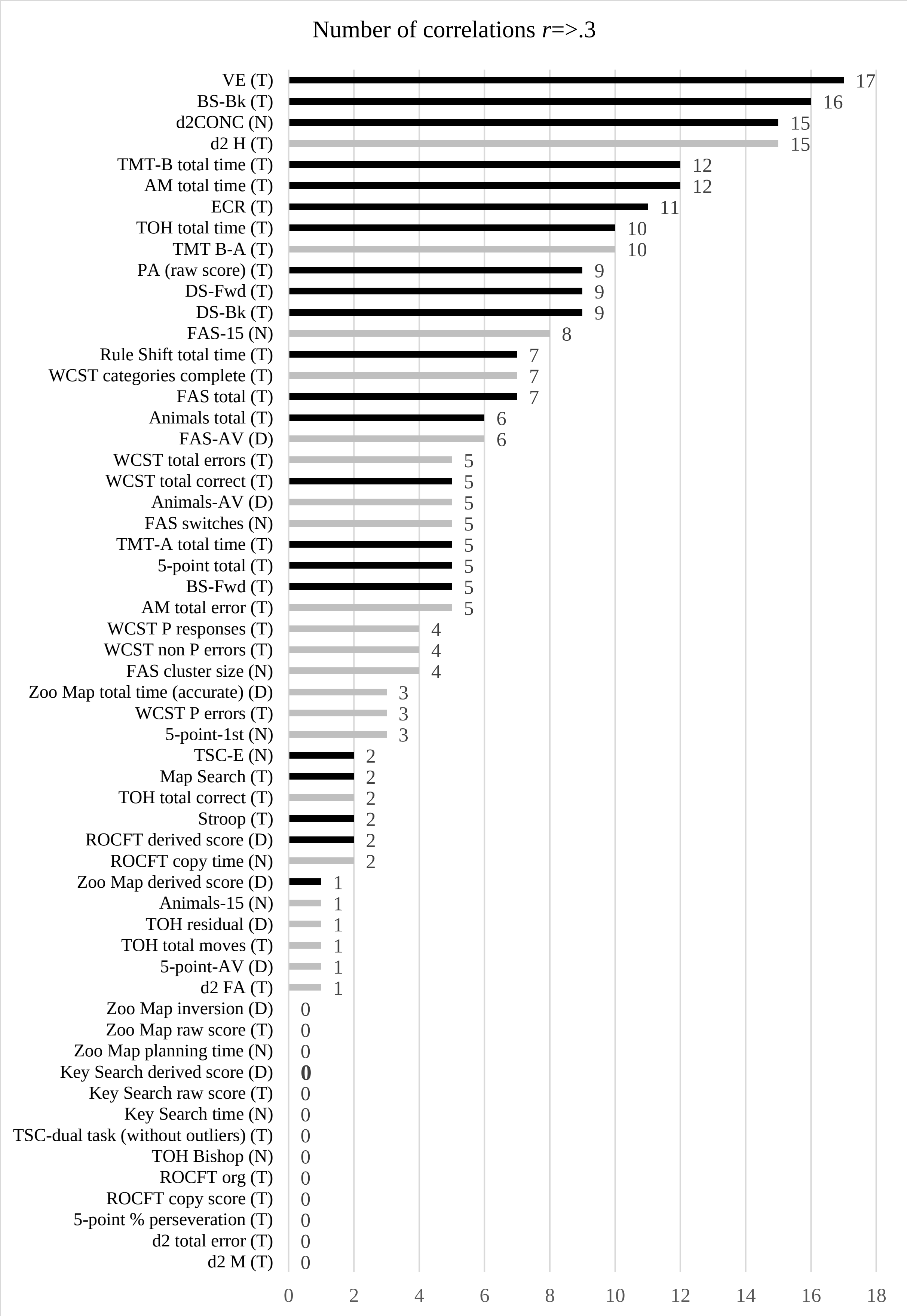


Figure 9. Frequency distribution of all 57 variables during the variable selection and

reduction phase.

Note: Boldface indicates retained variable. (T)=Traditional scoring approach. (N)=Non-traditional scoring. (D)=Derived: scoring procedures created by researcher.

Figure 9 above depicts each variable and the magnitude of significant correlations yielded in descending order. It was clear that most often traditional scoring approaches demonstrated a greater number of correlations compared to non-traditional scoring, or derived scores. Of note, despite some tests lacking clear variance and number of correlations above .3 they were retained because they were supported by theory, thereby satisfying criterion 2b for test inclusion. These tests were the ROCFT derived score, Zoo Map derived score, and Key Search derived score. A table of test inclusion as driven by Anderson where possible, or variations of tests can be viewed in Tables 4-7 below. In summary, of the 57 variables available, the data reduction method through means of Pearson's correlation yielded 23 variables appropriate for further analyses.

The following section outlines the placement of the retained tests in relation to Anderson's model of EF. A rationale for the tests used in the current study to represent each construct of his model including Attentional Control, Cognitive Flexibility, Information Processing and Goal Setting will be put forward.

Table 4

Test Inclusion to Confirm Anderson's Attentional Control Construct

Anderson's theory	Anderson's purported tests	Present study's test inclusion to confirm Anderson's AC construct
Selective attention, self-regulation, self-monitoring, inhibition	A not B task, Day-Night Test, Shape School Test, Conflict Task, TMT	BS-Fwd[a] (T) BS-Bk[b] (T) DS-Fwd[c] (T) DS-Bk[d] (T) TMT-A[e] (T) Map Search (T)

[a=]Block Span Forwards [b=]Block Span Backwards [c=]Digit Span Forwards [d=]Digit Span Backwards [e=]Trail Making Test-A

As indicated in Table 4 above, all span tasks have been chosen on the basis of them reflecting an attention and concentration component, as well some aspects of WM as described within this domain. These could be argued to exert low level demand on WM capacity, as participants do not require the manipulation of stimuli for these tests to be considered in the Cognitive Flexibility domain. Moreover, the TMT-A was chosen Anderson (2001) suggests this task is useful at detecting self-regulation, impulsivity and selective attention problems. Lastly, Map Search was chosen as this is purported to measure selective attention.

Table 5

Test Inclusion to Confirm Anderson's Cognitive Flexibility Construct

Anderson's theory	Anderson's purported tests	Present study's test inclusion to confirm Anderson's CF construct
Divided attention, WM[a], conceptual transfer, feedback utilisation	Stroop, WCST[b], Concept Generation, Contingency Naming Test	Stroop (T) WCST total correct (T) VE[c] (T) ECR[d] (T) TSC-E[e] (N) TMT-B[f] (T)

[a]=Working Memory [b]=Wisconsin Card Sorting Test [c]=Visual Elevator [d]=Elevator Counting with Reversal [e]=Telephone Search while Counting-E (does not attempt to control for the decrement and individual variation in processing or psychomotor speed) [f]=Trail Making Test-B

As indicated in Table 5 above, the WCST and the Stroop test have been placed in keeping with Anderson's model. Many of the TEA tests were chosen as these represented a divided attention aspect, which also encompasses WM. Furthermore, although not explicitly placed within this sub domain, Anderson (2001) proposes that TMT-B is useful in identifying divided attention and mental flexibility for middle childhood (6-12 years), and given this model is being validated in an adult population and not in a paediatric sample, the addition of this task is warranted.

Whilst the WCST did not deviate from traditional scoring approaches, this did however elicit multiple outcomes measures according to scoring manuals. Closer inspection revealed that these measures were highly correlated to one another almost to the point of multicollinearity, as expected with several measures used from a single test. Therefore, total correct was selected. This also allowed avoidance of the violation of the assumption of observation of independence. For example, perseverative and non-perseverative errors are the combination summed to create total errors, and total errors is the residual of total correct, meaning that the same observation was used for multiple scores, hence why total correct was only used in the current study.

Table 6

Test Inclusion to Confirm Anderson's Information Processing Construct

Anderson's theory	Anderson's purported tests	Present study's test inclusion to confirm Anderson's IP construct
Efficiency, fluency, speed of processing	COWAT[a]	FAS total[b] (T)
		Animals total[c] (T)
		5-point total (T)
		Rule Shift (T)
		d2CONC[d] (N)

[a=]Controlled Oral Word Association Test [b=]Verbal Fluency phonemic category FAS [c=]Verbal Fluency semantic category animals [d=]d2 test of attention concentration score

As indicated in Table 6 above, the Verbal Fluency tests (FAS and Animals) have been placed as per Anderson's description. 5-point total has also been placed here as a measure of non-verbal fluency. Rule Shift has been placed here as this could be argued to be a speed of processing skill. All of the tests under this subdomain are timed tests or are under time constraints to reflect a speed of processing skill, which as described previously entails both fluency and efficiency to achieve an overall 'better' score.

An example of the value of test and scoring consideration is offered by the purported relationship between strategy utilisation and Verbal Fluency, whereby scores were micro-analysed non-traditionally according to Troyer et al. (1997) observing clustering and switching strategies, and according to Delis et al. (2001) for task initiation and maintenance through each 15 second interval. The same break down principle of scoring was also applied to the non-verbal equivalent of the 5-point test. Study findings indicated that generally speaking the same pattern of correlations were evident across all fluency tasks, where the strength and number of correlations were greater for total correct compared to the break-down of scoring. This could be because the current study computed an average of the last intervals (last three for Verbal Fluency, last two for non-Verbal Fluency) for both measures.

Similarly, the test of d2 was scored in a non-traditional way (d2CONC) to measure both speed and accuracy equally according to Bates and Lemay (2004). However, study findings indicated that whilst the non-traditional concentrate score (d2CONC) demonstrated the same number of correlations to the traditional scoring of d2 H, the strength of correlations for the concentrate score was superior, suggesting that in this instance, the non-traditional approach to scoring was more beneficial.

Table 7

Test Inclusion to Confirm Anderson's Goal Setting Construct

Anderson's theory	Anderson's purported tests	Present study's test inclusion to confirm Anderson's GS construct
Initiative, conceptual reasoning, planning, strategic organisation	ROCFT[a], Porteus Maze, Everyday Problem Solving Inventory, Strategy Application Test, Cognitive Estimation, TOL[b]	ROCFT derived score (D) AM total time[c] (T) Zoo Map derived score (D) Key Search derived score (D) PA[d] (T) TOH total time[e] (T)

[a]=Rey Osterrieth Complex Figure Task [b]=Tower of London [c]=Austin Maze [d]=Picture Arrangement [e]=Tower of Hanoi

As indicated in Table 7 above, similar tests (or variants of the test) have been chosen to Anderson's. For example, TOH, AM, and ROCFT have all been placed here to represent planning and organisation. However, initiate is difficult to measure given this is more or less a qualitative measure. The Austin Maze (AM) was chosen as Anderson also used a maze task, and Zoo Map and Key Search have been placed here as these represent a planning and organisational component that fits within Anderson's sub domain.

In general, study findings indicated that measures relating to time were most often retained according to the number of correlations compared to those that utilised total correct, number of errors, or number of moves. Whilst in some instances total errors or total correct measures still demonstrated a number of correlations, it was evident that total time was far superior.

It is clear from the review of the literature that the TOH is a task that engages multiple skills with varying degrees of performance. Thus, multiple outcome measures were considered for the present study. The amount of errors on the TOH was not considered as a variable because this violated assumptions of normality, and therefore five different scoring protocols were considered including number of moves, total time, total trials correct (correct was determined if they completed in the minimum number of moves), residual (how many more moves over the minimum it took to complete) and a scoring approach described by Bishop (2001). However, study findings indicated that total time demonstrated five times more relationships to other tests compared to the remaining scoring protocols. This same pattern was evident within the Austin Maze. Results indicated that AM total time demonstrated more relationships to other measures than did errors. Whilst errors are central to performance on the maze and the error score still demonstrated variability, it was evident that time produced double the amount of relationships.

Study findings indicated that the ROCFT copy and organisation scores did not correlate with other measures. This is consistent with research that found no differences between the organisation score between younger and older adults (Gagnon, Awad, Mertens, & Messier, 2010). As reported throughout Strauss et al. (2006) it is clear that the copy score is able to differentiate different disorders. However, as evident in current study findings it is clear that the copy score is not as useful in healthy samples. Thus, whilst the copy score did demonstrate variability, its usefulness in the current context was limited as most participants obtained close to the maximum score possible. The total copy time demonstrated variability and two relationships to other measures, as did the derived score. Thus, the derived score was retained that used a combination of both speed and accuracy.

The same pattern was evident for measures of the BADS, in particular Key Search and Zoo Map. Of the measures of the Zoo Map, only total time and the derived score demonstrated correlations to other measures. Whilst planning time demonstrated variability, it seems that total time was superior given the outcome of the correlation results. Furthermore, the raw score demonstrated lack of variability, because participants in the study did not make many errors on this task. After consideration, although total time demonstrated three correlations, the derived score that included a combined speed and accuracy score that only demonstrated one relationship above the threshold set was retained. The use of this score addressed the lack of variability within the raw score, thereby utilising a speed and accuracy score as this test was required to confirm Anderson's Goal Setting construct. Lastly, the Key Search variables did not demonstrate any correlations to other measures, and for the same reasons above, the derived score was also retained to address lack of variability from the raw score, as this was also required to confirm the Goal Setting construct. Thus, in the instances where traditional, non-traditional, or deriving both speed and accuracy score did not provide value, scores were derived to create an overall efficiency score. This occurred for three tests: the ROCFT, Zoo Map, and Key Search.

The most likely explanation for the patterns reported in the data is because of the healthy adult sample used. Healthy adults rarely make many errors on EF tasks, therefore error scores lacked variability within a test, and were not valuable. This premise was particularly evident on tasks designed from deficit models such as those from the BADS tests. Therefore, it could be suggested that total time provides more value in distinguishing performance between individuals from healthy samples, whereas performance based on errors are of more value clinically

This same pattern of performance is consistent with previous research that found Zoo Map planning time failed to produce correlations clinically, whereas raw scores that included errors did (Oosterman, Wijers, & Kessels, 2012). Similarly, Bennett et al.'s (2005) study using a clinical sample found that both the Key Search and Zoo Map (assuming raw scores as this is standard scoring) made it to their final factor solution. This therefore provides support for total correct or total errors being valuable clinically.

In contrast, Testa et al.'s (2012) healthy sample failed to provide support for the Key Search task (raw score on test 1) in their factor analytic study. The authors used the raw score which relies on errors, and their sample almost obtained the maximum 14 points (*M*= 12.48, SD= 3.53), consistent with the current findings, albeit on test 2 (raw score; *M*=12.89, SD= 3.58), indicating a lack of variability. This therefore provides support that a healthy sample should not rely on total correct or errors.

Finally, whilst this study has found that a limited number of correlations were evident for tasks demanding more complex cognitive performance (e.g., Goal Setting), from the number of correlations it was evident that tasks representing elements of attention demonstrated an increased number of relationships to span tasks, VE, ECR, d2 H, d2CONC and trails measures. Thus, EF measures that theoretically are considered indicative of less complex cognition demonstrated a greater number of correlations with other variables than variables that were more complex. Whether this conclusion is a matter of mathematical artefact is unlikely, as the advantage of this study was the comprehensive battery completed by *all* participants. Furthermore, all tests were administered in counterbalanced order to ensure that waning attention was somewhat controlled.

Given that tasks with an attentional component, or lower down on a hypothesised continuum of complexity demonstrated the greatest numbers of correlations, it could be suggested that attention is fundamental to all cognitive performance. This is consistent with Sarter et al. (2001) who postulate "attention represents a basic attentional function that determines the efficacy of the 'higher' aspects of attention (selective attention, divided attention) and of cognitive capacity in general" (p. 147).

To this end given that many of the tasks that failed to produce meaningful relationships to other tests were most commonly tests of planning or Goal Setting, these findings may also indicate the difficulties associated with measuring tasks considered "higher order" or more complex. This was evident even when measures of time that have proved valuable in a healthy sample failed to produce many correlations in tasks considered higher order. A possible reason could be that theoretically, any higher order EF test can measure a number of separate, yet related EF skills that are inherently difficult to assess in isolation, consistent with the task impurity problem (Miyake et al., 2000). This may explain why statistically, attempting to break down scores to isolate a measure was not as strong, compared to total scores when other skills are combined. Thus, it may be that the variance shared amongst complex tasks is not captured by a smaller sample size, with lots of sub skills subsumed by effective task performance.

In summary, contrary to expectation, 'traditional' scoring approaches offered more utility than alternative scoring, as represented by a larger number, and sometimes strength of correlations across a variety of EF domains as demonstrated in Figure 9 above. For the clinician and researcher alike, test selection is a contentious issue that can greatly bias results. For the clinician, it is the balance between an appropriate number of measures against time constraints imposed in a clinical setting. For the researcher, emphasising too many tests or measures with a narrow focus will impact outcomes. Thus, if a researcher were to use the same approach as a clinician, this would limit the ability to find enough variables to represent the construct being measured. Thus, it is essential that researchers undertake a comprehensive approach to task selection because the outcomes of their endeavours inform clinical practice.

Study findings indicated that EF tests can be measured in different ways (e,g., speed and errors), and each measure often provided multiple outcome scores with varying utility across latent constructs. This is consistent with previous literature suggesting that any given EF test can measure a number of separate, yet related EF skills. However, this thesis was constrained by using some tests that were not designed for healthy populations (e.g., measures of the BADS), and therefore these findings suggest that in a healthy adult sample, measures relating to total errors or total correct are not appropriate, and instead total time would be more appropriate because this creates variability within tests. This thesis has argued that in order to keep purity in our tests, it is vital to consider both measures of speed and accuracy together relative to the task at hand, however study findings demonstrated that this was difficult to achieve. Whilst time has proven a vital measure providing additional value to both the research and clinical communities, future research should focus on developing tests appropriately aimed at healthy populations because clearly tests designed from a deficit model are not sensitive enough to distinguish performance between individuals in healthy samples. Thus, developing a model of EF in a healthy population is not only beneficial, but essential for the clinical community, so that inferences can be made on what atypical performance should present itself as. Nonetheless, time was a valuable measure because time necessarily inflates when errors are made. Therefore, in a way, speed and 'accuracy' were somewhat captured by time alone. However, what a measure of time alone is not able to do is distinguish between those whose disposition leads them to slow down to accommodate for increased task demands (i.e., speed-accuracy trade-off), and those who have lower capacity for information processing.

Interestingly, it is also clear from these findings that as a task becomes more complex, in particular tasks that require planning and organisation, the ability to find meaningful relationships amongst them becomes relatively obscured, consistent with EF theory. These are some of the reasons why correlation analyses that were designed to reduce the number of variables for inclusion for subsequent analyses fell short, particularly for tasks considered higher order. Findings indicated that of the vast number of tests available to clinicians, a smaller battery of tests can be utilised for efficient assessment of EF. The implications of these findings call for future researchers to shift away from using tests designed using a deficit model in healthy adult samples, and instead develop a more appropriate test battery for healthy adults whereby time is more important over errors or total correct, particularly for model development.

7.3 Data Preparation for The Confirmation of Anderson's Purported Latent Constructs in a Healthy Adult Population Using Congeneric Modelling

This section describes the meticulous approach taken to data preparation for the confirmation of Anderson's model using the SEM approach of congeneric modelling. The previous section outlined that of the 57 variables included in the current study, the variable selection and reduction method yielded 23 variables to be used for the congeneric models (that will use a CFA approach). Assumptions specific to SEM are described below. Of note, linearity, homoscedasticity and independence of observations were upheld following appropriate screening for multivariate assumptions.

7.3.1 Missing data

There are several challenges that exist when dealing with missing data in SEM, and careful consideration was given to the best solution for the different measures. Considering the importance of retaining variables for theoretical purposes and to avoid specification error, due to the limitations highlighted above regarding all the WCST variables, the selected key variable WCST total correct from the variable selection and reduction phase (7.5% missing) was submitted to MVA this time and revealed Little's MCAR test was also non-significant Chi-square=000, df= 201, p=1.000. This left a total of N=133 participants for all constructs to test except for Goal Setting, because of the amount of data missing for the TOH and AM variables as described above. However, given SEM requires at least three indicators per construct, it was essential to retain certain variables. Inspection revealed that although there were differences in the amount of data missing for each separately, there were 20 cases in which both TOH total time and AM total time were missing equally for each test. As such these cases were deleted due to the amount of missing data, leaving TOH and AM at 8.8% and .9% missing respectively. Little's MCAR revealed a desired non-significant result Chi square= 19.748, df=206, p=1.000 and the remaining values were then replaced via

Expectation Maximisation. This left a remaining *N*=113 for confirming the Goal Setting construct, and the other constructs remained at *N*=133.

7.3.2 Normality

SEM is sensitive to deviations from Multivariate Normality (MVN) however, this is often impractical to examine (Kline, 2015) as restrictions imposed by raw empirical data rarely achieve this assumption (Gao, Mokhtarian, & Johnston, 2008). Thus, meeting univariate normality on every variable often suffices (Kline, 2015). This method is also tolerated because Maximum Likelihood Estimation (MLE) as the estimation method for SEM is particularly robust to violations of normality (Hu & Bentler, 1999). Univariate normality was checked using AMOS and no values fell outside the acceptable cut off points of 3 for skewness and 10 for kurtosis (Kline, 2015). MVN was checked as a precaution through Mardia's coefficient and the critical ratio (c.r) was greater than 1.96 on all models indicating potential threats to MVN (largest value was 5.402). Since the c.r exceeded 1.96, bootstrapped bias corrected confidence intervals (C.I) were calculated to compare the *p* values with those generated without the bootstrapped standard error. The Attentional Control (AC) model indicated that TMT-A was a significant indicator prior to bootstrapping, however returned a non-significant value. The standard error bias however was small (.048) indicating an acceptable fit, however results regarding this test must be interpreted with caution. Furthermore, Map Search was a non-significant predictor before and after bootstrapping, indicating concerns of multivariate non-normality and should also be interpreted with caution. The remaining models (IP, CF, GS) indicated consistent *p* values after bootstrapping and small biases. It should be noted that because the TOH and AM were the only variables with a pure measure of time in the GS model, this resulted in an increase in variability compared to the other tests in the GS model which were derived (combining both speed and accuracy), thus making the variability smaller. Thus, the downside of this is an increased C. I, and the bias slightly larger (approximately -2.5). Furthermore, when no modification indices were evident Key Search returned a non-significant value, however when modification indices were demonstrated Key Search remained a significant indictor before and after bootstrapping. Therefore, given the consistent comparable *p* values and small biases, the present data can be confidently declared to be free from threats of multivariate non-normality.

7.3.3 Outliers

As explained above univariate outliers were treated according to various methods, where only extreme values were treated, meaning those not considered extreme were left untreated. This was done given the smaller sample size of the study as it was not feasible to continue to delete cases. This was also done in order to obtain a true representation of the data (Orr et al., 1991), because deleting non-extreme outliers would result in a loss of observations, which is obviously undesirable (Gao et al., 2008). However, as previously discussed the MLE is particularly robust to minor violations and the means for calculating an acceptable cut off in the literature is scarce (Gao et al., 2008). Mahalanobis distance was also checked and met in SPSS for all models except for Goal Setting (Mahalanobis critical value= 22.46, sample value= 24.16), where Cook's D was .09. However, this is close to the acceptable cut off of 1 (Allen & Bennett, 2012; Tabachnick & Fidell, 2013), and most likely occurred as a result of the limited non-extreme outliers, and therefore should be interpreted with caution. However, because they did not exceed 1 this indicates that the influence of the data did not hinder the predictive efficacy of the model as a whole (Allen & Bennett, 2012; Judd & McClelland, 1989).

7.3.4 Sample size

As previously explained, the total sample size was 133 participants (for all constructs apart from Goal Setting where N=113). SEM is a large sample technique and much research has focused on appropriate sample sizes for SEM (Bentler & Yuan, 1999; Hair et al., 1998;

MacCallum et al., 1996), and recommendations range between 100-200 (Hair et al., 1998). However, there is no consensus on what constitutes an appropriate sample size (Jackson, 2003), and considering some estimation models have been developed for as few as 60 participants despite some limitations (Bentler & Yuan, 1999), our sample seems reasonable. Hair et al. (1998) and Raykov and Widaman (1995) suggest there are a number of factors that impact upon, and therefore determine, adequate sample size. Based on their descriptions, *model misspecification* did not occur because a comprehensive approach was applied to include all theoretically relevant variables (as much as possible) and therefore this was not a problem that should be applied to the current data. Furthermore, rules surrounding *model size* suggests a minimum of 5 respondents per parameter, with 10 per parameter being adequate (Bentler & Chou, 1987; Hair et al., 1998), and some have suggested as little as 3:1 or 2:1 (Bagozzi and Yi, 2012). All models apart from Goal Setting did not exceed the 10:1 guidelines, therefore the respondent to parameter ratio meets this assumption. However, Goal Setting required 13 parameters and with the reduction in sample size to N=113, the minimum of 5 per indicator meets this assumption (13x5=65). There were no extreme *departures from normality*, therefore an increase of 15 respondents per parameter was not needed, and therefore Maximum Likelihood Estimation could be used as the *estimation procedure*, which assumes multivariate normality, and therefore allows valid interpretations for smaller samples of 100-150 (Hair et al., 1998). Strong factor loadings have also been known to mitigate any issues concerning small sample size (Wolf et al., 2013), however current findings did not demonstrate such strength. Nonetheless, for the reasons outlined above the current study’s sample of 133 seems reasonable as it is supported by the parameters detailed above, however caution is advised when interpreting findings.

7.3.5 Overall model fit

Multiple measures of fit were used as suggested by researchers (Hair et al., 1998; 2014; Hu & Bentler, 1998; Kline, 2015), particularly since modifications were made to each model as is acceptable when building models for theoretical purposes (Bagozzi & Yi, 2012). Each model was assessed for its validity in accordance to the following guidelines, where it is generally accepted to report a range of fit indices (McDonald & Ho, 2002). These are summarised in Table 8 below.

Table 8

Absolute Fit Indices

Fit Indices	Description	Acceptable cut off
GFI	Goodness of Fit Index	0-.95 indicates acceptable fit
AGFI	Adjusted Goodness of Fit Index	>.95 indicates acceptable fit
SRMR	Standardised Root Mean Square Residual	Lower the better indicates a better fit, however others suggest less than .08 will suffice
NFI	Normed Fit Index	Higher the better indicates a better fit, however others suggest >.95
IFI	Incremental Fit Index	Higher the better indicates a better fit, however others suggest >.95
CFI	Comparative Fit Index	Higher the better indicates a better fit, however others suggest >.95
χ^2	Discrepancy chi Square	Non-significant value is desired for an acceptable fit

RMSEA	Root Mean Square Error Approximation	Lower the better indicates a better fit, however others suggest less than .06 will suffice

The summary of several types of fit in Table 8 above was drawn from Hair et al. (2014), Kline (2015) and Hu and Bentler (1999). *Absolute Fit Indices* are measured on the basis of having no other models for comparison, and the most commonly reported is the chi-square statistic (χ^2) indicating the 'badness of fit' where a non-significant value is desired to demonstrate an acceptable fit (Hair et al., 2014). However, this statistic is rather sensitive to sample and model size. Thus, GFI (Goodness-of-fit index) and AGFI (Adjusted Goodness-of-fit index) were also used, where values between .90-.95 indicates acceptable fit (Hair et al., 2014; Hoelter, 1983 as cited in Hair et al., 2014). However, these often have the same issues as the χ^2 as they are simple transformations, and GFI does not perform as well with latent variables. SRMR (Standardised Root Mean Square Residual) was also used where a smaller value indicates better fit (Hair et al., 2014) with Hu and Bentler, (1999), suggesting less than .08. *Relative Fit Indices* (or comparative fit) are measured on the basis of a null-model or baseline measure where all measures are ‘uncorrelated’, where a desired larger chi-square value of .90 (Hair et al., 2014) to .95 (Hu & Bentler, 1999) for this null model indicates a poor fit. NFI (Normed Fit Index), IFI (Incremental Fit Index) and CFI (Comparative Fit Index) were also used, since CFI is considered appropriate for smaller sample sizes (Hu & Bentler, 1999; Kline, 2015). *Non centrality-based Indices* are measured on the basis of rejecting the alternative hypothesis (H1) (meaning why a smaller value is desired), which is in contrast to the chi-square which tests the acceptance of the null hypothesis (H0) (meaning why a larger value is desired). RMSEA (Root Mean Square Error Approximation) was used where lower values indicate a better fit (Hair et al., 2014), and are at least .06, although a lower value is desired (Hu & Bentler, 1999). Moreover, in order to test if a construct is in fact congeneric, models were compared to "parallel measures" (Lord & Novick, 1968), which assumes all measures are equal in their contribution towards a latent construct. Thus, parallel models were compared ('nested') to congeneric models which still assumes all measures are a true reflection of a construct, yet contribute differently. Therefore, to test if a congeneric model is a better fitting model than a parallel model, comparisons between the two were examined by comparing the χ^2 value of each, including which of the two was more significant, and which was a better fitting model. Lastly, standardised residuals should be centred around 0, and must not exceed 1.96 (Anderson & Gerbing, 1988). Essentially, larger numbers indicate a better fit for all indices reported except for RMSEA and SRMR.

7.3.6 Measurement model fit: reliability, validity and composite scores

In order for a model to be considered congeneric or unidimensional there are a few rules of thumb. However, there are discrepancies noted in the literature as to how many factors to retain, what the cut of values are, what overall fit indices to consider, and in particular, which of the many reliability estimates is best (Bacon, Sauer, & Young, 1995). For example, reliability measures are used to assess validity, and vice versa. Therefore, in the instances where estimates of reliability and validity were required to assess whether the indicators were sufficient in their representation of the constructs (Hair et al., 1998), the following reliability and validity formulas were calculated, with various cut off values used as a guide.

An assumption of CFA is that a construct must be unidimensional. However, this assumption in psychological research is difficult to obtain given that measures are often complex (Widhiarso & Ravand, 2014), and the assumption does not always hold (Brunner & Süβ, 2005 as cited in Widhiarso & Ravand, 2014). Thus, psychological constructs are rarely unidimensional (Slocum-Gori, Zumbo, Michalos, & Diener, 2009) and unidimensionality does not always infer that psychological data only measure one process (Bejar, 1983, as cited in Ziegler & Hagemann, 2015). Thus, if differences within a latent construct are due to different psychological processes, it is often ensured that each test should adequately reflect these processes (Fischer, 1997 as cited in Ziegler & Hagemann, 2015). This is particularly useful in relation to the present findings given the multidimensional aspects of EF. Therefore, using the incorrect coefficient can be harmful, particularly if values decrease causing negative patterns (e.g. when lower scores for time indicate better performance). This is not only important for the reliability and validity of the scale, but also when computing a factor score that takes into consideration the unique weighting of each indicator, which is the premise of a congeneric model. Thus, the term composite reliability (also referred to as construct reliability) refers to the reliability of composite scores, and it is important to consider the correct formula based on these caveats. Therefore, the hierarchical omega coefficient (also known as coefficient H) was used (Hancock & Mueller 2001), which is considered to be a maximised optimum construct reliability measure, for which a negative sign does not impact the overall assessment of reliability, unlike standard methods (e.g., Raykov, 1997). This formula performs well on multidimensional measures by weighting each indicator based on factor loadings (Widhiarso, & Ravand, 2014). Coefficient H was used in conjunction with traditional construct reliability and validity measures (Average Variance Explained), because many reliability measures are used to assess validity, and vice versa. The formulas and guidelines are expressed below;

Construct/composite reliability according to coefficient H

$$H = \frac{1}{\left[1 + \left[\frac{1}{\frac{\lambda_1^2}{1-\lambda_1^2} + \frac{\lambda_2^2}{1-\lambda_2^2} + \dots + \frac{\lambda_n^2}{1-\lambda_n^2}}\right]\right]}$$

where the λ's are the standardised factor loadings.

However, **construct validity** is made up of multiple components (Hair et al., 2014), and therefore other formulas were considered. For example, **convergent validity** assesses the proportion of variance variables have in common, in that they should converge to represent an overall construct. This is assessed by the size of the factor loadings, where acceptable cut-offs range from .5-.8 (Bagozzi, 1991; Fornell & Larcker, 1981; Hair et al., 2014; Nunnally & Bernstein, 1994 as cited in Hancock Mueller, 2001). Furthermore, an individual reliability of an indicator should be approximately .5, which roughly equates to a standardised loading of .7 (Hair et al., 1998), however others have suggested as low as .4 (Zainudin, 2012). All indicators should also be significantly different from zero, however Bagozzi and Yi (2012) suggest that hypotheses and goodness-of-fit should be the emphasis, as often, indicators may be low, with satisfactory performance. Convergent validity is also assessed using the average variance extracted formula below. Thus, reliability and validity analyses should be interpreted with caution, with the stated caveats in mind.

Construct/composite reliability according to Average Variance Extracted (AVE) reflects the overall amount of variance in the indicators accounted for by the latent construct (Fornell & Larcker, 1981). The AVE is proposed by Fornell and Larcker (1981), where the reliability for each measure and AVE can be calculated in a similar fashion. It is recommended that the AVE should exceed at least .5 (Fornell & Larker, 1981; Hair et al., 1998), and is presented below:

$$\rho_{vc(\eta)} = \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \sum \theta_i}$$

where λ_i is the standardised loading for each observed variable and θ_i is the error variance associated with each observed variable.

Discriminant validity is demonstrated when the model is free of redundancy (i.e., no modification indices suggest covariance), however researchers have suggested allowing the covariance of error values only with theoretical justification (Bagozzi & Yi, 1998 as cited in Zhang, 2015).

Composite scores were calculated according to Jöreskog and Sörbom (1989) after fitting and accepting a one-factor congeneric model. This takes into account factor score weights of each indicator variable, variance of the factor, estimated (or implied) covariance matrix and error variances. Factor score weights are proportional to the factor loadings and error variances so as to avoid incorrect estimation if simple unit weight addition of scores were to be used.

7.4 Data Preparation for Assessing the Predictive Validity and Interrelationships of Verified Constructs in a Healthy Adult Population Using Regression Analyses

7.4.1 Missing data, linearity, homoscedasticity, multicollinearity, normality, outliers

Composite scores were calculated based on reliable and valid constructs that were identified through congeneric modelling in the previous step. All assumptions to satisfy standard multiple regression analyses were conducted and met, however it was evident that there was too much missing data for the Goal Setting construct (15% missing, *N*=113) to be comparable to the remaining constructs (AC, CF, IP, *N*=133). It was therefore necessary to rectify in order to confirm the interrelationships. Thus, the same deleted 20 ID cases from the TOH and AM were deleted from the AC, CF, and IP constructs, leaving a total sample of *N*= 113 for the remaining analyses making all constructs comparable. These methods were the most appropriate given the restricted sample size.

Chapter 8 Results

8.1 Confirmation of Anderson's Purported Latent Constructs in a Healthy Adult Population Using Congeneric Modelling

Data reduction processes yielded 23 variables for testing the confirmation of the four constructs in Anderson's model in a healthy adult population. A minimum requirement for this analysis would be 12 variables therefore the inclusion of 23 is a strength of the study as the requirements of a CFA approach is that at least three reflective indicators per construct were needed. Table 9 below displays the means, standard deviation, z scores, and range of performance across the variables selected.

Table 9

Means, Standard Deviations, Raw scores, Z scores, and Possible Range of Performance of Variables Selected

Variable	N	M *(SD)*	Minimum value *(Z score)*	Maximum value *(Z score)*	Range
Attentional Control					
BS-Fwd[a]	133	6.77 *(1.91)*	2.00 *(-2.49)*	11.00 *(2.21)*	0-16
BS-Bk[b]	133	6.55 (1.84)	3.00 *(-1.92)*	12.00 *(2.96)*	0-16
DS-Fwd[c]	133	6.74 *(2.31)*	2.00 *(-2.06)*	14.00 *(3.14)*	0-16
DS-Bk[d]	133	6.08 *(2.66)*	1.00 *(-1.91)*	14.00 *(2.98)*	0-16
TMT-A[e]	133	23.22 *(6.91)*	10.56 *(-1.83)*	49.81 *(3.85)*	0-180
Map Search	133	68.72 *(8.65)*	43.00 *(-2.97)*	80.00 *(1.30)*	0-80
Information Processing					
d2CONC[f]	133	185.12 *(44.49)*	28.00 *(-3.53)*	294.00 *(2.45)*	-58-300
Animals total[g]	133	26.62 *(5.31)*	13.00 *(-2.56)*	41.00 *(2.71)*	0+
FAS total[h]	133	42.42 *(10.92)*	16.00 *(-2.42)*	71.00 *(2.62)*	0+
5-point total[i]	133	34.57 *(8.03)*	9.00 *(-3.18)*	53.00 *(2.30)*	0+
Rule Shift[j]	133	27.94 *(6.78)*	15.26 *(-1.87)*	56.50 *(4.21)*	0-∞

Cognitive Flexibility					
WCST total correct[k]	133	48.43 *(9.37)*	15.00 *(-3.57)*	59.00 *(1.13)*	0-64
VE[l]	133	3.77 *(0.86)*	2.00 *(-2.06)*	7.30 *(4.10)*	-
ECR[m]	133	7.90 *(2.23)*	2.00 *(-2.64)*	10.00 *(0.94)*	0-10
TSC-E[n]	133	2.95 *(0.88)*	1.70 *(-1.41)*	5.30 *(2.66)*	-
Stroop[o]	133	1.68 *(0.43)*	0.99 *(-1.63)*	2.83 *(2.67)*	0-∞
TMT-B[p]	133	57.18 *(17.80)*	20.06 *(-2.08)*	114.50 *(3.22)*	0-300
Goal Setting					
Zoo Map derived score[q]	113	7.37 *(6.60)*	0.85 *(-0.99)*	23.60 *(2.46)*	-
Key Search derived score[r]	113	5.50 *(4.55)*	0.54 *(-1.09)*	16.43 *(2.40)*	-
ROCFT derived score[s]	113	4.52 *(1.51)*	2.17 *(-1.55)*	11.17 *(4.40)*	-
TOH total time[t]	113	312.37 *(118.65)*	76.85 *(-1.98)*	727.80 *(3.50)*	0-∞
AM total time[u]	113	305.25 *(79.33)*	143.34 *(-2.04)*	483.07 *(2.24)*	0-∞
PA[v]	113	13.40 *(3.66)*	4.00 *(-2.57)*	19.00 *(1.53)*	0-22

[a=]Block Span Forward trials correct [b=]Block Span Backwards trials correct [c=]Digit Span Forward trials correct [d=]Digit Span Backwards trials correct [e=]Trail Making Test-A total time to complete [f=]d2 concentrate Hits-False Alarms [g=]Verbal Fluency semantic total admissible words in 3 minutes [h=]Verbal Fluency phonemic total admissible words in 3 minutes [i=]5-point total admissible designs in 3 minutes [j=]Rule Shift total time to complete [k=]Wisconsin Card Sorting Task total correct responses [l=]Visual Elevator timing score [m=]Elevator Counting with Reversal total correct [n=]Telephone Search while Counting weighted for accuracy of tone counting (measure of divided attention) [o=]Stroop test interference score [p=]Trail Making Test part B total time to complete [q=]Zoo Map 2 total time divided by total correct [r=]Key Search time divided by total correct [s=]Rey Osterrieth Complex Figure Task time divided by total correct [t=]Tower of Hanoi total time to complete all 13 trials [u=]Austin Maze total time to complete (trials to criterion; 2 error free trials) [v=]Picture Arrangement total raw score.
Note: Maximum ranges for Fluency measures (Animals, FAS, 5-point) have not been provided because outcome scores depend upon a time limit rather than a set possible score. Similarly, the Stroop, TOH, and AM have no maximum score as they are also timed. Lastly, the derived scores (TSC-E, Zoo Map, Key Search, ROCFT) do not have a range because these were computed for the purpose of the present study and can also vary based on completion time.

8.1.1 Attentional Control

A congeneric model was created to test the underlying structure of the latent construct Attentional Control. Results indicated that, BS-Fwd, BS-Bk, DS-Fwd, DS-Bk, Map Search and TMT-A were overall not good fitting measures of Attentional Control. Initial analysis revealed a poor fitting model was evident χ^2= 40.82 with 9 *df* (*p* <.001). RMSEA= .16, AGFI= .77, GFI= .90, CFI= .78, IFI= .78, NFI= .74 which is most likely a result from the non-significant indicator Map Search (*p*=.349). Thus, Map Search was deleted and analysis revealed that a poor fitting model was still evident χ^2 =17.35 with 5 *df* (*p*= .004), RMSEA= .14, AGFI= .85, GFI= .95, CFI= .90, IFI= .90, NFI= .87. Modification indices revealed that the error variances for BS-Fwd and BS-Bk needed to be covaried to improve the model (M.I= 7.160, Par Change=.706). Similarly, BS-Bk and TMT-A also needed to be covaried (M.I= 5.759, Par Change= -2.417), and therefore BS-Bk was removed. A good fitting model was then identified χ^2= 1.61 with 2 *df* (*p*= .447), RMSEA= .00, SRMR= .039 AGFI= .97, GFI= .99, CFI= 1.00, IFI= 1.00, NFI= .98 demonstrating a good model fit. The standardised residuals were no greater than 1.96 also indicating a good fit, in addition to all indicator variables being significant at the .001 level except for TMT-A (*p*=.044), and all error variances being significant at the .001 level except for DS-Bk (*p*=.034) and DS-Fwd (*p*=.003). A parallel analysis was conducted to ensure the model was in fact congeneric, and results revealed the congeneric model was a significantly better model than the parallel, Chi-square 398.131 with 6 *df*, *p* <.001. Overall, the good fit of the model suggested that BS-Fwd, DS-Bk, DS-Fwd and TMT-A were good measures of a single underlying construct of Attentional Control. A further break down of measures is presented below. Figure 10 displays the path analysis.

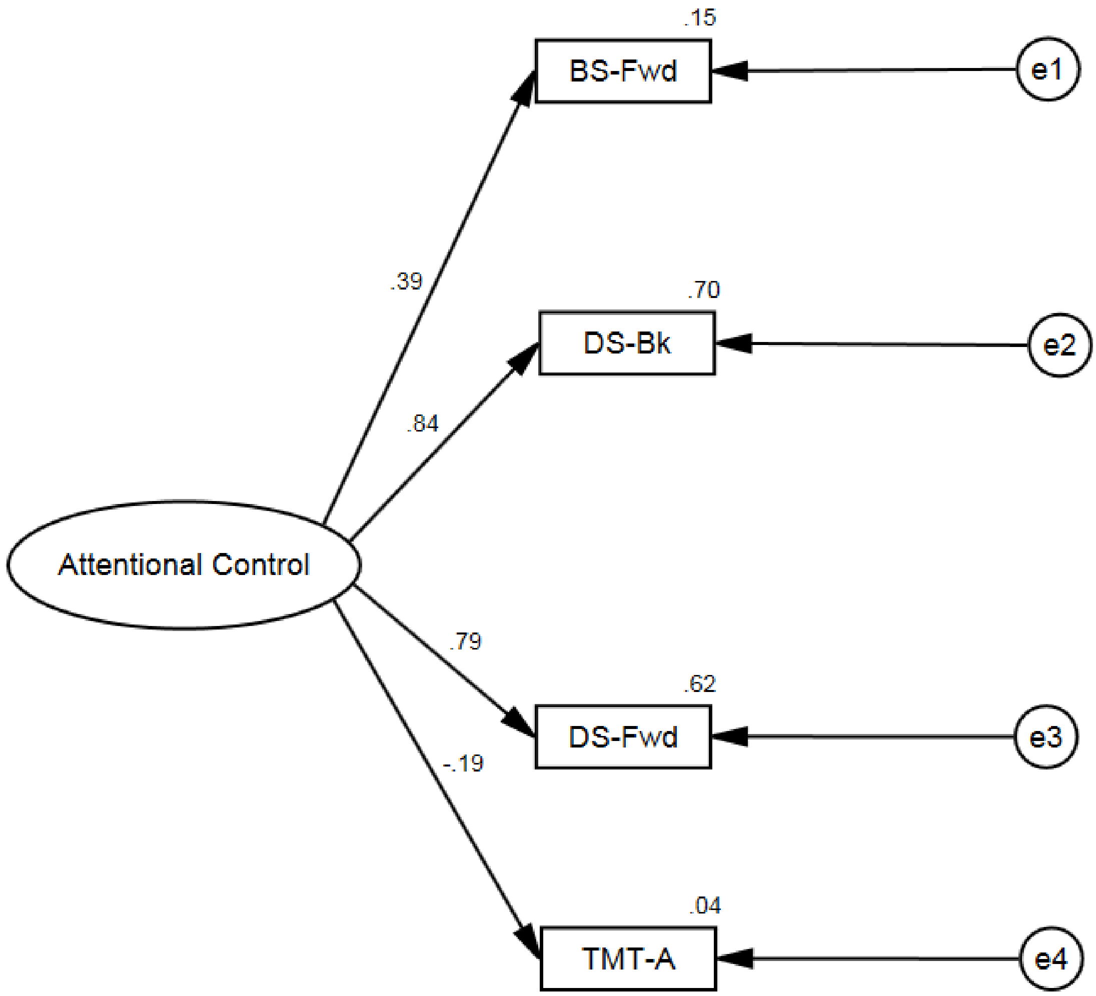


Figure 10. Path diagram of Attentional Control *N*=133.

As indicated in Figure 10 above, the squared multiple correlations (SMC) indicate the reliability of the indicator variables and demonstrated that the latent construct Attentional Control accounts for between 4 and 70% of the indicators. These SMC reliabilities reported in Figure 10 above mostly meet the recommended limit of .5 with the exception of BS-Fwd and TMT-A, which were still below the absolute minimum of .4. All regression factor loadings and error variances were greater than the critical value 1.96 demonstrating significance, whereby two of the four factor loadings were above the recommended limit of

.5-.7. The exception to this were BS-Fwd and TMT-A (as Attentional Control goes up by one, TMT-A goes down by .19 standard deviations).

The AVE coefficient was .377 which is under the cut off of .5, thus more than half the variance for these indicators is not accounted for by the construct of Attentional Control. Although some individual loadings are acceptable (except for BS-Fwd and TMT-A) the variance explained falls short. However, construct reliability was evident since coefficient H returned a value of .806 indicating stronger maximised reliability, and discriminant validity was demonstrated as the model was free of redundancy.

In summary, these variables form a significant congeneric measure of the Attentional Control latent construct indicated by the overall model fit and construct reliability coefficients, however individual factor loadings and reliabilities should be interpreted with caution on an individual variable basis.

8.1.2 Cognitive Flexibility

A congeneric model was created to test the underlying structure of the latent construct Cognitive Flexibility. Results indicated that WCST total correct, VE, ECR, TSC-E, the Stroop and TMT-B were good fitting measures of Cognitive Flexibility. Initial analysis revealed that a good fitting model was evident χ^2 = 13.31 with 9 *df* (*p*= .149), RMSEA= .06, AGFI= .92, GFI= .97, CFI= .97, IFI= .97, NFI= .91, however modification indices suggested the error variance of TSC-E and Stroop needed to be covaried to improve the model (M.I.= 9.696, Par Change .085). A better fitting model was then identified χ^2 = 2.901 with 8 *df* (*p*= .940), RMSEA= .000, SRMR= .026, AGFI= .98, GFI= .99, CFI= 1.00, IFI= 1.04, NFI= .98, demonstrating a better model fit. The standardised residuals were no greater than 1.96 also indicating a good fit. All indicator variables were significant at the .001 level and all error variances were significant at the .001 level. A parallel analysis was conducted to ensure the model was in fact congeneric, and results revealed the congeneric model was a significantly better model than the parallel chi-square 1731.875 with 10 *df*, *p* <.001. The good fit of the model suggests that WCST total correct, VE, ECR, TSC-E, Stroop and TMT-B are good measures of a single underlying construct Cognitive Flexibility. Figure11 below displays the path analysis, and a further break down of measures is presented below.

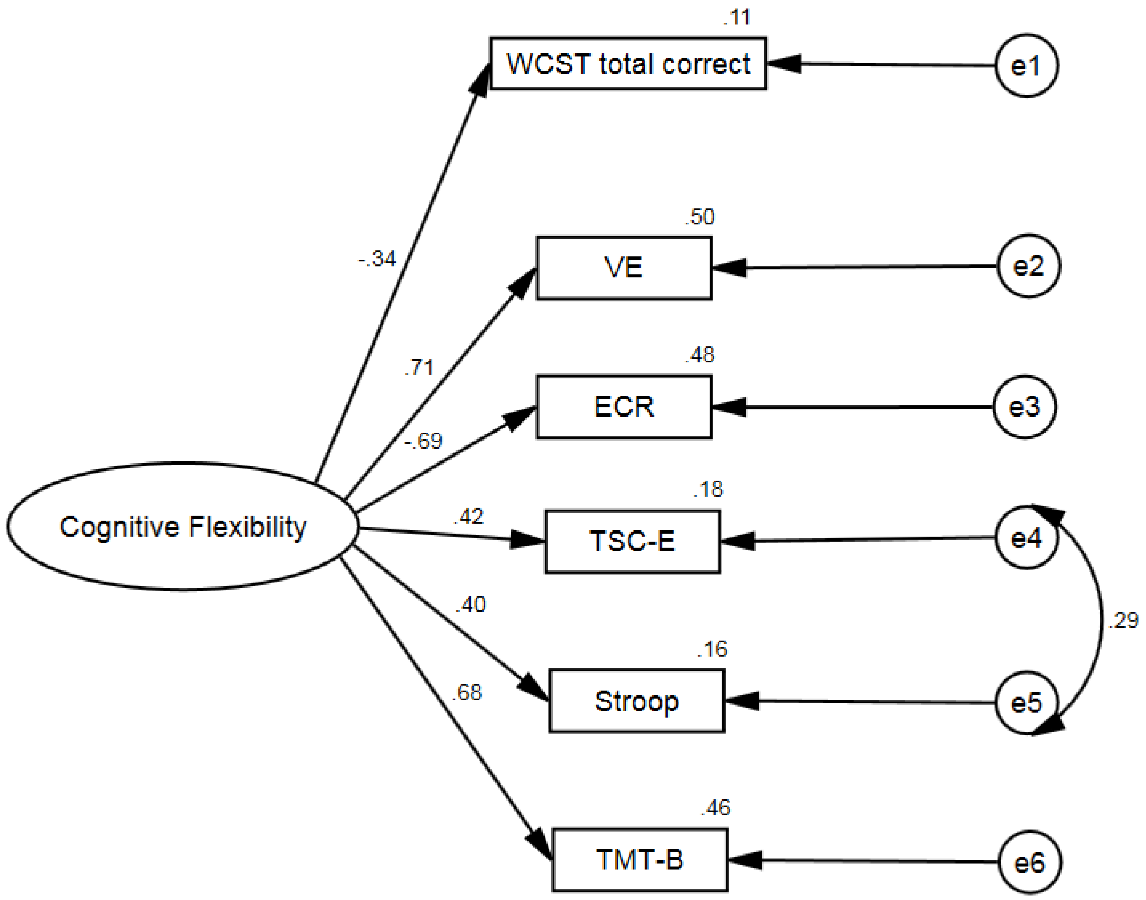


Figure 11. Path diagram of Cognitive Flexibility *N*=133.

As indicated in Figure11 above, the squared multiple correlations (SMC) indicate the reliability of the indicator variables, and demonstrated that the latent construct Cognitive Flexibility accounts for between 11 and 50% of the indicators. These SMC reliabilities reported in Figure 11 above almost meet the recommended limit of .5 with the exception of

WCST total correct, TSC-E, and the Stroop test which were below the absolute minimum of .4. All regression factor loadings and error variances were greater than the critical value 1.96 demonstrating significance, whereby three of the six factor loadings were above the recommended limit of .5-.7. The exception to this were WCST total correct, TSC-E, and the Stroop (as Cognitive Flexibility goes up by one, the Stroop goes up by .40 standard deviations).

The AVE, the coefficient was .315 which is under the cut off of .5, thus more than half the variance for these indicators is not accounted for by the construct of Cognitive Flexibility. Although some individual loadings are acceptable (except for WCST total correct, TSC-E, and Stroop) the variance explained falls short. However, construct reliability was evident since coefficient H returned a value of .767 indicating stronger maximised reliability. Discriminant validity however was only somewhat demonstrated, and it could be argued on theoretical grounds that TSC-E and the Stroop test were covaried. This could be based on the premise of a shifting skill required for both tasks. It is also argued that they both require high level task demand in that for the Stroop one must inhibit a strong pre potent response, and although individually the demands of the TSC-E are simple, when combined they require one to concentrate and complete two tasks simultaneously. The correlation between the two was $r= .29$ indicating a weak positive relationship.

In summary, these variables form a significant congeneric measure of the Cognitive Flexibility latent construct indicated by the overall model fit and construct reliability coefficients, although individual factor loadings and reliabilities should be interpreted with caution on an individual variable basis.

8.1.3 Information Processing

A congeneric model was created to test the underlying structure of the latent construct Information Processing. Results indicated that d2CONC, Animals total, FAS total, 5-point total and Rule Shift time were overall good fitting measures of Information Processing. Initial analysis revealed that an arguably good fitting model was evident $\chi^2 = 7.68$ with 5 *df* (*p*= .175), RMSEA= .06, AGFI= .93, GFI= .98 CFI= .96, IFI= .96, NFI= .90, however modification indices revealed there was need for the error variance of Animals total and FAS total to be covaried to improve the model (M.I. 4.619, Par Change 9.068). A better fitting model was then identified $\chi^2 = 1.15$ with 4 *df* (*p*= .886), RMSEA= .000, SRMR= .022, AGFI= .99, GFI= 1.00, CFI= 1.00, IFI= 1.04, NFI= .98 demonstrating a better model fit. The standardised residuals were no greater than 1.96 also indicating a good fit, in addition to all indicator variables and error variances being significant at the .001 level. A parallel analysis was conducted to ensure the model was in fact congeneric, and results revealed the congeneric model was a significantly better model than the parallel chi-square 865.090 with 8 *df*, *p* <.001. Overall, the good fit of the model suggests that d2CONC, Animals total, FAS total, 5-point total and Rule Shift time are good measures of a single underlying latent construct Information Processing. Figure 12 below displays the path analysis and a further break down of measures are presented below.

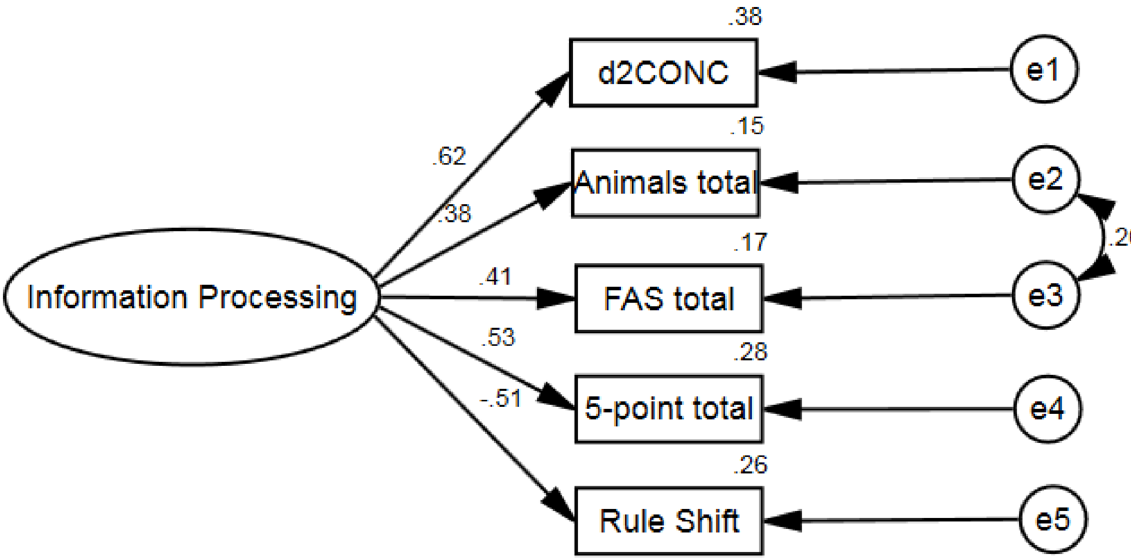


Figure 12. Path diagram of Information Processing *N*=133.

As indicated in Figure 12 above, the squared multiple correlations (SMC) indicate the reliability of the indicator variables and demonstrated that the latent construct Information Processing accounts for between 15 and 38% of the indicators. Of these SMC reliabilities reported in Figure 12 above, d2CONC could arguably be said to meet the absolute minimum of .4 (.38), where Animals total, FAS total, 5-point total, and Rule Shift were below the cut off. All regression factor loadings and errors variances were greater than the critical value 1.96 indicating significance, whereby three of the five factor loadings were above the recommended limit of .5-.7. The exception to this were Animals and FAS total (as Information Processing goes up by one, Animals total goes up by .38 standard deviations).

The AVE coefficient was .247 which is under the cut off of .5, thus more than half the variance for these indicators is not accounted for by the construct of Information Processing. Although some individual loadings are acceptable (except for Animals and FAS total) the variance explained falls short. However, construct reliability was evident since coefficient H returned a value of .633 indicating stronger maximised reliability. Discriminant validity is cautiously assumed, however theoretical reasons suggest that covarying Animals and FAS is not a concern and is expected as they require the same skills to be used in a different way with a correlation of *r*= .26 indicating a weak positive relationship.

In summary, these variables form a significant congeneric measure of the Information Processing latent construct indicated by the overall model fit and construct reliability coefficients, although individual factor loadings and reliabilities should be interpreted with caution on an individual variable basis.

8.1.4 Goal Setting

A congeneric model was created to investigate the unidimensional underlying skills of the latent construct Goal Setting. Results indicated that a moderate fitting model was evident χ^2 = 13.93 with 9 *df* (*p*=.125), RMSEA= .07, AGFI= .90, GFI= .96, CFI= .90, IFI= .91, NFI= .78 however modification indices suggested the need for the error variances of TOH total time and PA to be covaried to improve the model (M.I= 4.25, Par Change= -77.868). A better fitting model was then identified χ^2 =7.22 with 8 *df*, (*p*=.513), RMSEA= .000, SRMR= .056, AGFI= .94, GFI= .98, CFI= 1.00, IFI=1.01, NFI=.88.The standardised residuals were no greater than 1.96 however TOH total time and Key Search derived score were -1.38, and ROCFT derived score and Zoo Map derived score were 1.42 which are quite close. Furthermore, all indicator variables were significant at p <.02. All error variances were also significant at the .001 level except for AM (*p*=.166). A parallel analysis was conducted to ensure the model was in fact congeneric, and results revealed the congeneric model was a significantly better model than the parallel chi-square 2316.04 with 10 *df*, p <.001.

Overall the good fit of the model suggests that the Zoo Map derived score, Key Search derived score, ROCFT derived score, TOH total time, AM total time and PA are good measures of a single underlying construct of Goal Setting. Figure 13 below displays the path analysis, and a further break down of measures are presented below.

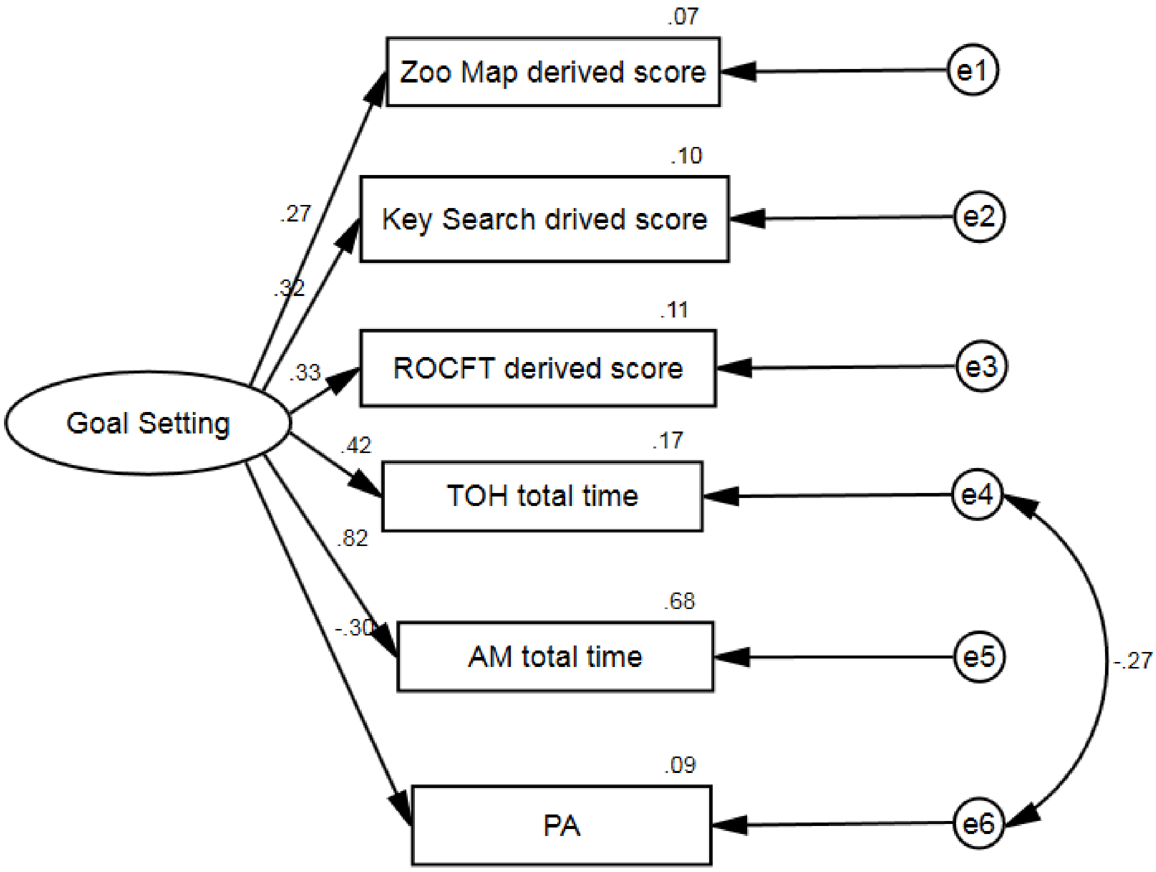


Figure 13. Path diagram of Goal Setting *N*=113.

As indicated in Figure 13 above, the squared multiple correlations (SMC) indicate the reliability of the indicator variables and demonstrated that the latent construct Goal Setting accounts for between 7 and 68% of the indicators. Of these SMC reliabilities reported in Figure 13 above, AM was the only indicator meeting the recommended limit of .5, where Zoo Map, Key Search, ROCFT, TOH and PA were well below the absolute minimum of .4. All regression factor loadings and error variances were greater than the critical value 1.96 indicating significance, except for the AM error variance which was not significant (p=.166), where AM was the only indicator above the recommended limit of .5-.7 (as Goal Setting goes up by one, AM total time goes up by .82 standard deviations).

The AVE coefficient was .203 which is under the cut off of .5, thus more than half the variance for these indicators is not accounted for by the construct of Goal Setting. Although only the AM indicator was acceptable and the variance explained falls short, construct validity was still demonstrated indicated by the overall model fit and since coefficient H returned a value of .730 indicating stronger maximised reliability. Discriminant validity is cautiously assumed, with the covarying of TOH time and PA. This could perhaps indicate non-verbal anticipation, organisation and planning. The correlation between the two was $r = -.27$ indicating a weak negative relationship.

In summary, these variables form a significant congeneric measure of the Goal Setting latent construct indicated by the overall model fit and construct reliability coefficients, although individual factor loadings and reliabilities should be interpreted with caution on an individual variable basis.

8.2 Assessing the Predictive Validity and Interrelationships of Verified Constructs in a Healthy Adult Population Using Regression Analyses

Composite scores were calculated using only those variables included in each of the final congeneric models. Composite scores were then used in a series of standard multiple regression analyses to determine the predictive validity and strength of interrelationships of each construct. Given the number of repeat analyses a more conservative alpha level was set to .01. Descriptive analyses and Pearson's correlation coefficients are presented in Table 10 below.

Table 10

Descriptive Data and Pearson's Correlation Coefficients of the Four Validated Constructs

Construct	N	M *(SD)*	Min	Max	IP	AC	CF
IP[a]	113	65.80 *(17.12)*	21.52	105.89			
AC[b]	113	6.00 *(1.91)*	1.54	11.55	.350**		
CF[c]	113	3.74 *(1.51)*	0.45	9.39	-.634**	-.424**	
GS[d]	113	36.14 *(9.35)*	20.27	59.81	-.484**	-.149	.475**

Note. ** correlation is significant at the .01 level (two-tailed).

[a=]Information Processing [b=]Attentional Control [c=]Cognitive Flexibility [d=]Goal Setting

As indicated in Table 10 above, all constructs demonstrated a moderate to strong significant relationship between each other. However, AC was found to be non-significant and weak to GS.

8.2.1 Cognitive Flexibility

A standard multiple regression was run to determine the predictive validity of AC, GS, and IP collectively on CF. A significant model was found F (3,109) = 34.52, p <.001, R=.698, *AdjustedR*^{2}= .473 indicating that 47.3% of the variance in CF, could be predicted by AC, GS, and IP. Table 11 presents this information below.

Table 11

Results of the Regression Analysis for Cognitive Flexibility

Construct	*B*	β	*t*	*p*	Partial *r*	sr^2	*R*	*AdjustedR*^{2}
AC[a]	-.186	-.235	-3.21	.002	-.294	4.8%		
GS[b]	.036	.226	2.88	.005	.266	3.8%		
IP[c]	-.039	-.442	-5.35	.000	-.456	13.4%		
Overall model							.698	.473

Note. Dependent variable = CF

[a=]Attentional Control [b=]Goal Setting [c=]Information Processing

As indicated in Table 11 above, all variables were significant predictors, and Information Processing holds the strongest relationship to Cognitive Flexibility.

8.2.2 Information Processing

A standard multiple regression was run to determine the predictive validity of AC, CF, and GS collectively on IP. A significant model was found F (3,109) = 30.43, p <.001, R=.675, *AdjustedR*^{2}= .441 indicating that 44.1% of the variance in IP, could be predicted by AC, CF, and GS. Table 12 presents this information below.

Table 12

Results of the Regression Analysis for Information Processing

Construct	*B*	β	*t*	*p*	Partial *r*	sr^2	*R*	*Adjusted*R^2
AC[a]	1.03	.115	1.47	.145	.139	1.0%		
CF[b]	-5.32	-.470	-5.35	.000	-.456	14.2%		
GS[c]	-.447	-.244	-3.03	.003	-.279	4.5%		
Overall model							.675	.441

Note. Dependent variable = IP

[a=]Attentional Control [b=]Cognitive Flexibility [c=]Goal Setting

As indicated in Table 12 above, CF and GS were both significant predictors to the model, however AC was non-significant. Cognitive Flexibility holds the strongest relationship to Information Processing.

8.2.3 Goal Setting

A standard multiple regression was run to determine the predictive validity of AC, CF, and IP collectively on GS. A significant model was found F (3,109) = 14.74, p <.001, R=.537, *Adjusted*R^2= .269 indicating that 26.9% of the variance in GS, could be predicted by AC, CF, and IP. Table 13 presents this information below.

Table 13

Results of the Regression Analysis for Goal Setting

Construct	*B*	β	*t*	*p*	Partial *r*	sr^2	*R*	*Adjusted*R^2
AC[a]	.467	.095	1.06	.290	.101	0.7%		
CF[b]	1.94	.313	2.88	.005	.266	5.3%		
IP[c]	-.174	-.319	-3.03	.003	-.279	6.0%		
Overall model							.537	.269

Note. Dependent variable = GS

[a=]Attentional Control [b=]Cognitive Flexibility [c=]Information Processing

As indicated in Table 13 above, both CF and IP were significant predictors, however AC was non-significant. Information Processing holds the strongest relationship to Goal Setting.

Overall results indicated that AC was not consistently statistically contributing to a model of EF. Therefore, the AC construct was removed from the model of EF, which therefore lent itself for further analyses.

A pictorial representation of regression analyses on constructs with subsumed tasks and skills can be viewed in Figure 14 below.

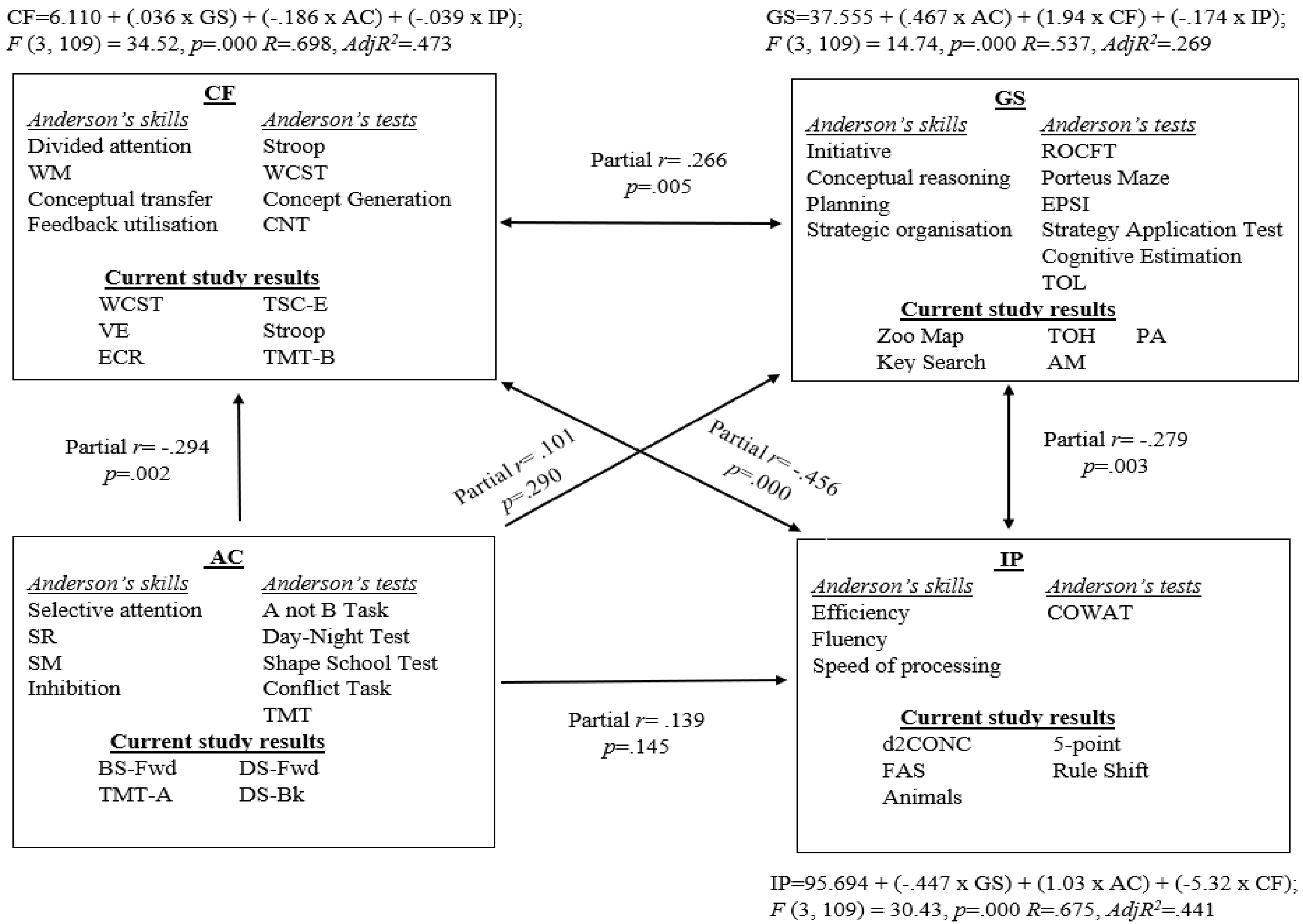


Figure 14. Pictorial representation of regression analyses on constructs with subsumed tasks N=113.

8.3 Investigating Information Processing as a Mediating Factor in Cognitive Performance

Results from the regression analyses indicated that the construct Attentional Control (AC) was a weak and non-significant predictor in Anderson's overall model of Executive Function. Thus, AC was removed which gave rise to a mediation analysis, to determine whether Information Processing (IP) (m) mediates the relationship between Cognitive Flexibility (CF) (x) and Goal Setting (GS) (y), given the strength of IP across all domains. Given the restricted sample size and after the removal of AC, the most appropriate parsimonious method to examine this relationship was through a mediation analysis.

The aim of the mediation analysis was to determine whether IP could explain the relationship between CF and GS. Statistical assumptions were the same as for the regression analyses, therefore all assumptions were met. Additional assumptions specified four conditions must be met in order for mediation to run. These are assumed to be tested through three regression models that identify the strengths of the relationships between variables and is outlined in Figure 15 below. Specifically;

1) The predictor variable must significantly predict the outcome variable *(denoted by path c)*
2) The predictor variable must significantly predict the mediator variable *(denoted by path a)*
3) The mediator variable must significantly predict the outcome variable *(denoted by path b)*
4) The predictor variable must predict the outcome variable significantly less compared to model 1, highlighting a reduction in the strength of the relationship once the mediator has been included, or reduced completely to zero *(denoted by c')* (Field, 2013).

The mediator is tested through bias-corrected and accelerated bootstrapping methods and achieves significance when the confidence intervals are entirely above zero after 5000 bootstrap resamples (Field, 2013). A proposed direct effect can be observed between CF and GS controlling for IP, and the indirect effect which is the effect of CF on GS through IP presented in Figure 15 below.

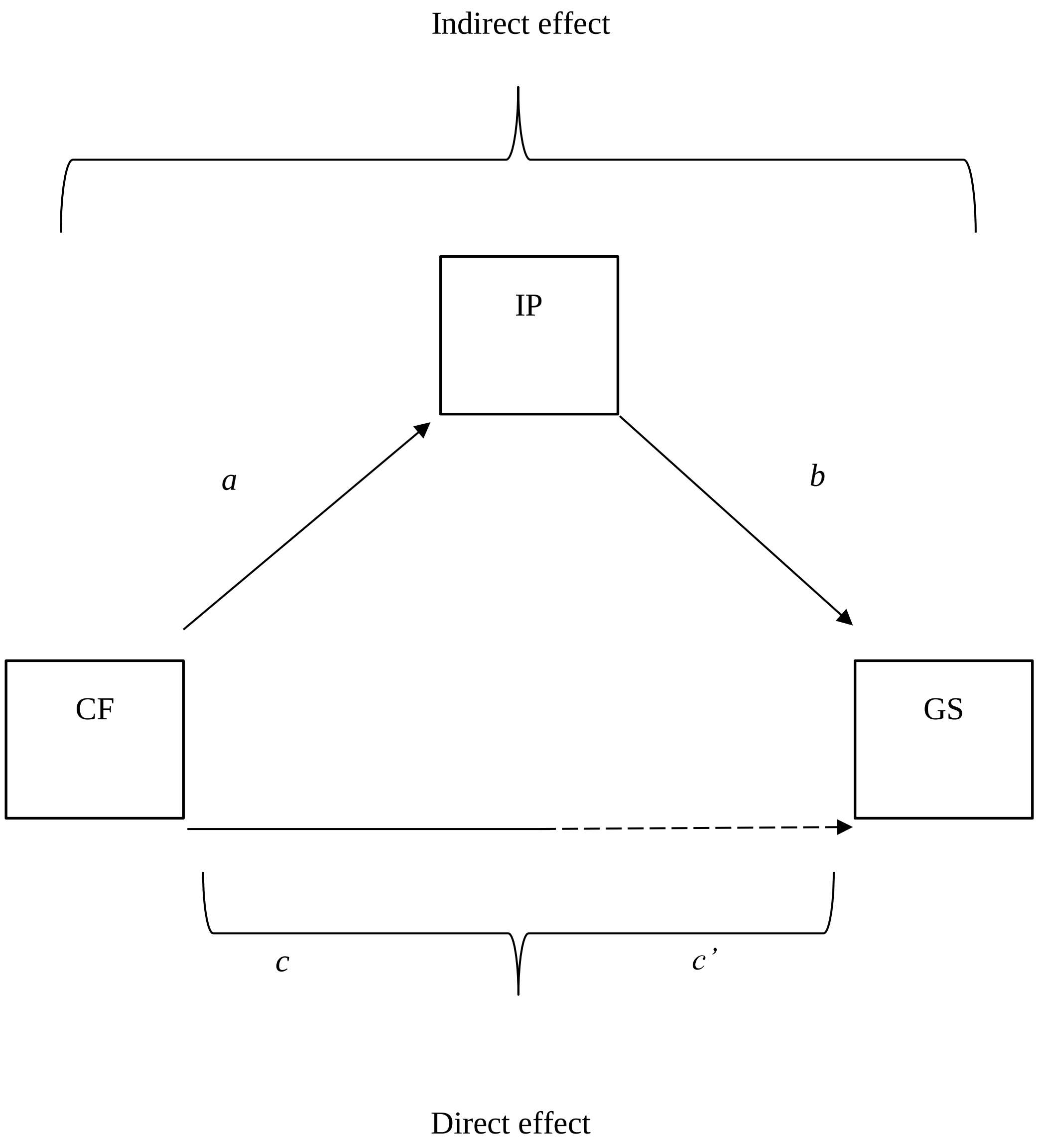


Figure 15. Proposed mediation model.

Results indicated (with α= .05) that Cognitive Flexibility was a significant predictor of Goal Setting F (1,111) = 32.27, p <.01, R^2= .23 accounting for 23% of the variance in GS. An increase in CF (where a higher score reflects better performance) by one unit lead to an increase in GS (where a higher score reflects better performance) by b=2.94, t= 5.68, p <.01 *(path c)*. Furthermore, Cognitive Flexibility was a significant predictor of Information Processing F (1,111) = 74.56, p <.01, R^2= .40, accounting for 40% of the variance in IP. An increase in CF by one unit lead to a decrease in IP (where a lower score reflects better performance) by b= -7.18, t=-8.64, p <.01 *(path a).* The model, which comprised of CF and IP, was a significant predictor of GS F (2, 110) = 21.52, p <.01, R^2 = .28 accounting for 28% of the variance in GS. It was found that IP was a significant predictor of GS when the effect of CF was held constant, where an increase in IP by one unit lead to a decrease in GS by b=-.17, t= -2.93, p <.01 *(path b).* The direct effect of CF and GS remained significant but weakened with IP in the model b= 1.74, t= 2.69, p=.01 *(path c')*. A bias-corrected and accelerated bootstrapped 95% confidence interval (CI) for the indirect effect was above zero, indicating a significant indirect effect of CF predicting GS through IP, consistent with partial mediation b= 1.20, 95% BCa CI [.40, 2.26]. This represented a medium effect ab_{cs} = .19, 95% BCa CI [.07, .35]. Figure 16 represents the partial mediation effect below.

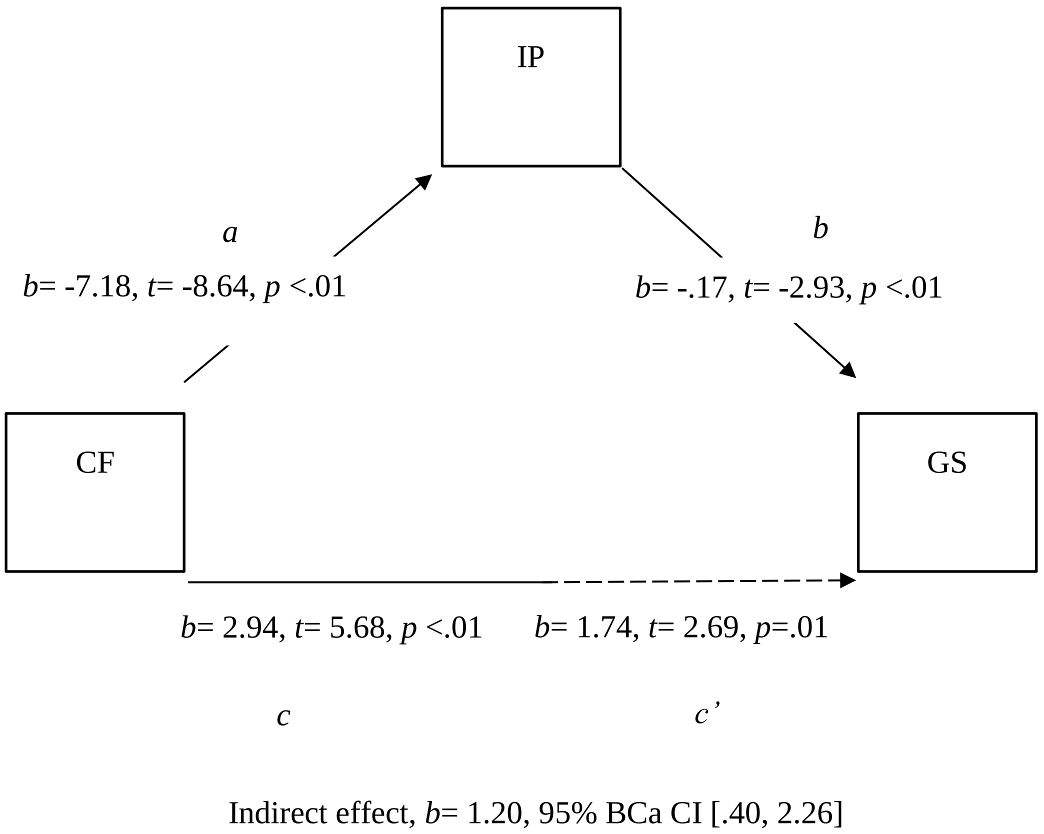


Figure 16. Indirect effect of Cognitive Flexibility and Goal Setting mediated by Information Processing. *N* = 113.

As demonstrated in Figure 16 above, results indicated that IP partially mediates the relationship between CF and GS. On its own, increases in CF led to higher GS, however when mediated through IP the strength of the relationship between CF and GS reduced.

Chapter 9 General Discussion

The overall research goal was to confirm the validity of the latent EF constructs purported by Anderson's model of EF in children, using a healthy adult sample. Furthermore, the research explored the strength of association and predictive validity between the identified constructs to achieve efficient goal directed behaviour. The overall methodological approach was guided by Testa et al. (2012), with the addition of further careful theoretical and statistical criteria applied to achieve a parsimonious yet comprehensive battery to explore the research questions.

The study findings overall, provide strong impetus for the assessment and validation of EF models using healthy adult samples. The benefits of this approach have been reflected in the way EF measures were shown to correlate, load and explain various specific and more global EF constructs. This study operationalised the mechanisms of EF through Anderson's model and it is argued that Information Processing is a major underpinning determinant of the efficiency of complex Executive Functioning. Whilst attention is a foundation skill arguably inherent in all aspects of cognition, it is the contention of this research that typical EF is more uniquely dependent on IP as operationalised in the current study. Given the developmental difference in population that Anderson's model of Executive Function was founded on (children) relative to the current study population (adults), it is further indicated that adults are not simply "big children", and that the role attention differs when EF skills are no longer emerging, but arguably established. Whilst this latter conclusion is intuitive, this research has significantly advanced the controversial and inconsistent inclusion of attention as synonymous with EF. Many of the definitional inconsistencies are addressed. Furthermore, by aligning the theoretical model with rigorous test selection, this research is able to offer a parsimonious battery of tests that reflect each of the underlining latent constructs effectively.

9.1 Confirmation of Anderson's Purported Latent Constructs in a Healthy Adult Sample Using Congeneric Modelling

The overall aim of the study was to confirm the validity of the latent constructs of Anderson's paediatric model of Executive Function in a healthy adult sample. Results indicated that all four constructs were statistically upheld and therefore "confirm" the validity of Anderson's paediatric model in healthy adults, thus supporting the first hypothesis.

Although statistical validation was indicated, there are numerous theoretical challenges that must be taken into account. Multiple facets contribute to the construct of Executive Functioning which is therefore considered multidimensional. Thus, it seems counterintuitive to test a one factor congeneric model built on the assumption of unidimensionality, and meeting this assumption required a careful balance of theoretical issues and definitions. This resulted in lower than recommended individual factor loadings, reliabilities, and statistical anomalies regarding directionality of tests (although theoretically sound), meaning the internal consistency of constructs should be interpreted with caution on an individual variable basis. This however, was not unexpected given that one EF task may measure multiple skills, but it also meant that some of the skills were not effectively assessed. In order to meet the minimum requirements for SEM, a congeneric model must have at least three indictors to represent a construct. Therefore, it was evident in the current findings that whilst it is optimal to have a parsimonious test battery to represent constructs, it was a difficult task in balancing between retaining enough variables to uphold construct validity (given the restrictions borne by multidimensional constructs) and removing ones that did not meet minimum loading requirements to avoid the dissolution of a construct in its entirety.

Furthermore, this thesis has identified there are a multitude of factors proposed by researchers that are thought to explain the construct of Executive Functioning. However, the varied scope of analyses including the methodology and samples used only hinder comparisons across studies and are in some way redundant. Thus, this thesis is constrained because direct comparisons are limited. Given the large healthy sample used, it is therefore sensible to compare large scale studies that utilise healthy adults such as Testa et al. (2012) and Miyake et al. (2000).

9.1.1 Differentiation between Attentional Control and Cognitive Flexibility

Discussing Attentional Control (AC) and Cognitive Flexibility (CF) as separate constructs in the current study proved challenging. Statistically, some of the tests placed under the AC construct demonstrated either i) lower than desired item reliabilities from an SEM perspective, suggesting they might not be consistently tapping into the construct. This was found despite evidence of good psychometric reliabilities of the tests overall during test selection. Or ii) did not demonstrate strong enough loadings and therefore were removed. From a theoretical viewpoint, this is in fact consistent with the multifactorial nature of EF and is likely why there was so much theoretical overlap between AC and CF. For example, Anderson described self- monitoring and regulation, inhibition, shifting, cognitive flexibility and WM are tapped into within the AC construct so that plans are executed in the correct order to ensure goals are reached. However, he also suggested that shifting, divided attention, WM, and attentional control such as self-regulation and monitoring, and inhibition are also tapped into when assessing the CF construct.

From the outset this thesis has acknowledged that Anderson's model of Executive Function is based on extensive literature and theoretical emphasis on the developmental trajectories in children. Whilst this thesis could not replicate Anderson's test utilisation exactly, every attempt was made to include measures that reflected the construct using adult normed tests. Despite this, considerable overlap was still evident, and therefore it is sensible to discuss both constructs in parallel.

9.1.2 Attentional Control

Attentional Control as a construct reflects self-regulation, monitoring, selective attention and inhibition, although with a healthy adult sample assessing these skills proved vulnerable to the limitations of traditional scoring. Tasks entered into Attentional Control were as close to the theoretical placement recognised by Anderson, such as span tasks, TMT-A, and Map Search. Map search was removed because this did not reach statistical significance which was unexpected because this typically is used as a measure of selective attention (Robertson et al., 1994). However even though not all measures were retained in the final model, a significant portion of variance of the overall construct was substantiated by the collective variables.

With respect to the construct of attention, further consideration was given to the measurement of an individual’s capacity for information retention. Span tasks offer the ability to measure an immediate span or capacity of information to be maintained (the famous 7 plus or minus 2 for verbal information), whilst the backward variant increases the task complexity. Thus, it expected that an individual would demonstrate a higher forward span than backward.

Study findings demonstrated that digits backwards, followed by digits forwards were the strongest indicators in the model. Block span backwards was removed because the error variance of this task was correlated with blocks forward and TMT-A, suggesting there may be common underlying mechanisms associated between all tasks. Furthermore, participants in the current study demonstrated a very similar average blocks forward and backward span, and the magnitude of difference between block forward and backwards was smaller than digits. Whilst it might be that blocks forward and backwards are equal in task demands, as suggested by previous research (Kessels, van den Berg, Ruis, & Brands, 2008), it may also be a reflection of measurement error covarying in the current study. TMT-A demonstrated poor strength of association to the AC construct, and the relationship it yielded between its error variance and block span backwards further contributes to the notion that measurement error contributed to the contrary results.

It was difficult to assess the sub skills of self-regulation and self-monitoring. Whilst it was an aim of the current study to provide a microanalysis of scoring systems, self-corrected errors presumed to reflect self-regulation and monitoring were scant, and therefore these skills were not effectively assessed. This most likely was because a healthy sample was used that were less likely to make sufficient errors. Therefore, in a healthy adult sample it can be assumed that both self-regulation and self-monitoring are fundamental to all cognitive performance across a variety of domains which may not be appropriately captured by error-based outcome scores. This is similar to Testa et al.'s (2012) study where they found a factor which they labelled Self-Monitoring and Set Maintenance which did not comprise errors, and instead, comprised the WCST (FB), RNG (repetition and cycling), and CNT (FC). The RNG test was used as a measure of repetitions of naming random numbers over 500 trials. This may be a direction for future research to use tests that do not rely on errors to tap into self-regulation and self-monitoring which may provide additional value when model development is in question, at least for healthy adults.

A limitation of this study was the lack of inhibition measures included. Likely owing to the awareness that inhibitory control is an emerging and maturing skill in children, inhibition may have been placed in either Attentional Control or Cognitive Flexibility. It was interesting to note that Testa et al.'s (2012) factor of Response Inhibition also did not include the Stroop, which, to the stated surprise of the authors, instead loaded onto a factor they labelled as Set Shifting and Interference Management. They argued that the Stroop test engages the "capacity to deal with more than one aspect of task stimuli concurrently rather than taxing their ability to inhibit a response" (p.220), and therefore requires one to "stay on task and cope with more than one task demand simultaneously" (p.220). Collectively previous study findings reinforced the placement of the lone inhibition measure employed in this study, the Stroop test, under the Cognitive Flexibility construct. Further research to clarify the similarities and differences in various inhibition tasks is critical, such as the difference types of inhibition highlighted by Barkley (1997) in his influential model of EF.

9.1.2.1 Reconceptualising simple and complex attention and capacity

Upon removal of inappropriate tests in the AC construct based on either statistical or theoretical grounds, it was evident that the remaining tests in the model best represent simple attentional capacity and is therefore reconceptualised as Simple Attentional Capacity (SAC). A possible reason why the tasks might be too simple to be accurately described as executive in nature could be because few cognitive resources are allocated to complete them. Thus, these findings could imply that the way information is processed depends in part upon the cognitive load required to complete it.

For example, selective attention theory models have proposed that stimuli are selected either early (Broadbent, 1958), intermediate (Treisman, 1964), or late (MacKay, 1973) in processing. High-load tasks use most, if not all of a person's resources thereby leaving no capacity to 'spill-over' to process and get distracted by irrelevant stimuli. For example, paying close attention to a particular task and being so 'tuned' in that distractions are unlikely to occur. In contrast, low-load tasks utilise few resources thereby leaving 'spare' capacity to 'spill-over' and process irrelevant stimuli, meaning distractions can occur. For example, not paying close attention to a task because of lack of interest leaves one vulnerable to distractions.

The tasks used were clearly "too simple" to be considered executive in a healthy adult sample. Thus, Anderson's conceptualisation of AC in children may represent complex control at lower stages of development, however for adults, it does not. This is because maturation of the frontal lobes follows a developmental trajectory through to early adulthood, therefore influencing the way in which skills develop to underpin this construct. In essence, the emphasis or maturity is different, and therefore AC tasks load differently between the two populations.

The single skills used under the SAC construct might explain why this represents a simple task load in healthy adults. For example, TMT-A is a visually oriented task (Crowe, 1998) with elements of speed of processing (Ríos, Periáñez, & Muñoz-Céspedes, 2004), and visual search skills (Mahurin et al., 2006), and Map Search has been purported as a measure of selective attention (Robertson et al., 1994). Given that the correlations from the variable selection and reduction phase demonstrated Map Search only correlated with TMT-A and TMT-B, it might suggest that both TMT-A and Map Search tasks would be facile for an unimpaired adult, and therefore do not fill up this cognitive load. TMT-A can be viewed as an automatic skill, because beyond visual search, task requirements are to essentially count from 1-13, and Map Search is simply searching for symbols. Therefore, because there are no other cognitive requirements concurrently demanded of these tasks, they are not classified as dual in nature, and are therefore classified as simply drawing on one skill.

Similarly, simple task load may also explain current findings regarding span tasks. Traditionally, backwards span tasks should be more difficult than forwards as they require more effort (Wechsler, 1997b), however in the current sample the capacity difference between forwards and backwards demonstrated a less than one-unit difference. Furthermore, proactive interference in the role of executive attention proposes that ‘simple’ attention tasks in the context of Working Memory Capacity (WMC) such as digit and block span, provide little clarity with respect to complex or ‘executive’ attention (Kane & Engle, 2003). Instead, they proposed that WMC has more insight than does attention alone and that complex WMC span tasks provide greater detail regarding executive attention. Thus, it may reflect the inattention of the sample because of the perceived perceptual load. To that end, on the grounds of Working Memory models (Baddeley, 1996) and Supervisory Attentional Control models requiring non-routine tasks, (Norman & Shallice, 1986), it is thought that this level of Attentional Control is mediated by the Central Executive, therefore facilitating WM to maintain goal relevant information (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002). Thus, it is no wonder that span tasks were not considered executive in nature, because if they can be performed using strategies such as chunking or rehearsal, then these could be implied to be considered 'automatic' thereby not taxing higher levels of Attentional Control akin to the SAS model, which therefore would not tax the Central Executive component of WM (Conway et al., 2002). It could be argued that Pires et al.'s (2018) CFA findings that demonstrated their span tasks that were used as a WM measure loaded onto a latent construct they labelled as EF, was only because their verbal span tasks could be considered more difficult. For example, every second item from a list was required to be recalled with a maximum of 16 digits, where scores could possibly range from 0-28. Thus, variability within their sets of scores was evident (mean trials correct in their study was 12.36 for digits, and 11.09 for blocks which is greater than the current study), where a composite of the two was considered as the final outcome measure.

Taken together, this might explain why the SAC construct explained the highest amount of shared variance. Effectively, a lower order test that represents a single skill would therefore share a higher proportion of variance to another lower order test as demonstrated in the variable selection and reduction phase. By contrast, anything that is considered higher order (or higher in the chain of processing than simple attentional capacity) would therefore draw on multiple skills, where one test is likely to reflect numerous skills and the individual contribution of each task would not be enough for a significant bidirectional relationship to occur. This would therefore explain the lower amount of shared variance in the CF construct.

The overlap between skills in the two constructs might be because whilst it still may be possible to assess some skills in isolation under a simple task load, it is clear that the CF construct explains the higher levels of Attentional Control that was lacking from the 'Simple Attentional Capacity' construct. In essence, distinguishing between the two constructs SAC and CF may come down to the dual nature of the task that increases in task complexity (e.g., two seemingly simple tasks required to be carried out simultaneously would therefore make the task more complex, and therefore share resources), where SAC reflects a bottom up approach, and CF a top down approach. Therefore, anything that is dual in nature should be considered higher or complex attention, and this is likely to be inherent in the Cognitive Flexibility construct. Therefore, complex attention should be used synonymously with Cognitive Flexibility, and therefore is considered a complexity bound construct.

9.1.3 Cognitive Flexibility

Cognitive Flexibility as a construct reflects divided attention, working memory, conceptual transfer, and feedback utilisation. Tasks entered into Cognitive Flexibility were more closely related to the theoretical placement recognised by Anderson compared to the AC construct, such as the WCST total correct, VE, ECR, TSC, Stroop, and TMT-B, and no tests were removed. Of the tests in this model the Cognitive Flexibility construct demonstrated the second highest amount of variance explained, and therefore this latent construct was confirmed in a healthy adult sample.

Specifically, study findings indicated that VE, ECR, and TMT-B were the strongest and most reliable tests in the model respectively. Notwithstanding the discrepancies in definitions in overall EF nomenclature, a review of the literature highlights that explanations regarding which skills may contribute to effective performance on cognitive flexibility tasks are lacking, and there are no models proposed to support this as a domain subsumed by a variety of skills. Therefore, together these tests strengthen the interpretation of the construct as being a measure of Cognitive Flexibility with its own set of attendant skills as driven by Anderson, because these tests are purporting to measure attentional switching and cognitive flexibility (Robertson et al., 1994) (VE), verbal WM (Robertson et al., 1994) (ECR), and set shifting, inhibition, attention control and cognitive flexibility (TMT-B) (Arbuthnott & Frank, 2000; Kortte, Horner, & Windham, 2002; Kowalczyk, McDonald, Cranney, & McMahon, 2001; Langenecker, Zubieta, Young, Akil, & Nielson, 2007; Ríos et al., 2004).

Study findings indicated that the TSC-E and the Stroop test were the weakest and most unreliable tests in the model, where the Stroop was the weaker of the two. This might be because the error variances were correlated to one another, which was also demonstrated in the variable selection and reduction phase where TSC-E only met a correlation above .3 with the Stroop. Given that traditionally the Telephone Search while Counting task is a measure of divided attention with elements of shifting or multitasking (Robertson et al., 1998), this might suggest that the Stroop test has elements of shifting to it. This is consistent with Testa et al.'s (2012) findings as explained in the SAC construct above, where they found the Stroop loaded onto a factor they labelled as Set Shifting and Interference Management, and argue the Stroop requires one to deal with more than one aspect of incoming stimuli at the same time (dual task). Testa et al.'s (2012) findings, however, were not consistent with Miyake and colleagues' (2000) who found that the Stroop was in fact related to an Inhibition factor. Instead, Miyake and colleagues argue that the Stroop is a prototypical measure of inhibition used to describe a variety of functions at different levels of complexity (Kok, 1999 as cited in Miyake et al., 2000). Thus, their interpretation of inhibition was restricted to a controlled, intended suppression of prepotent responses (as opposed to an automatic response in relation to negative activation or reactive inhibition in relation to negative priming). Furthermore, according to Barkley's (1997) placement of Behavioural Inhibition below in the chain of processing skills as a foundational prerequisite skill to higher order functions, it might be sensible to suggest that inhibition as represented by the Stroop should reconcile its placement with the SAC construct.

Lastly, study findings indicated the WCST total correct was the weakest and most unreliable indicator in the model. Traditionally, the WCST has been reported as a measure of cognitive flexibility so it is surprising that this was not a strong indicator. It might be because only total correct as a variable was used. This variable was selected for inclusion because all WCST variables were extremely correlated to each other, as expected with several measures used from a single test. Reasons why previous studies may have found support for a variety of WCST variables may be because of the statistical methods employed. Principle Components Analysis (PCA) partitions variance (Hair et al., 2014), and would therefore not be affected by multicollinearity. However, other latent approaches such as CFA do not partition variance, and instead, estimate shared and unique properties of the reliable variance (Strauss et al., 2000). Even still, combining several measures of the same test may be redundant particularly for the WCST. For example, perseverative and non-perseverative errors are the combination of total errors, and total errors is the residual of total correct, hence why total correct was the only score used in the current study. Therefore, it is unlikely that the purported skills conceptual transfer, or feedback utilisation were effectively assessed.

Consistent with previous definitions of cognitive flexibility suggesting incorporating multiple sources of information simultaneously as essential for this construct (e.g., Anderson 1998), the fact that VE and TMT-B that are described as measures of attentional switching, shifting, and cognitive flexibility that demonstrated the highest loadings and reliabilities under the CF construct, only strengthens the argument that anything dual in nature should be considered higher order. These tasks require one to maintain counting, whilst temporarily holding the last number in the mind and remembering the rules of the arrows to switch the direction of counting (VE), or to shift from numbers to letters (TMT-B). Furthermore, ECR as another strong indicator also strengthens the distinction explained above between span tasks being considered simple and WMC tasks considered more complex. ECR is a measure of verbal WM (Robertson et al., 1994) an arguably complex or higher order task. A key distinction between span tasks and the ECR is that ECR requires the manipulation of information for a short period of time, thus considered a WM task which is why the strength of this loading was evident. Therefore, clearly WMC span tasks such as the ECR task provide greater detail regarding 'executive attention' compared with simple attentional span tasks that do not require the manipulation of information. This is consistent with Barkley (1988) who suggested complex attention and Executive Functions often overlap and are difficult separate. Therefore, together these findings suggest that complex attention should be used synonymously with Cognitive Flexibility.

9.1.3.1 Distinguishing Attention and Cognitive Flexibility

Whilst this study has demonstrated there is a clear distinction between simple (essentially one skill) and complex (essentially dual nature of the task) tasks and constructs respectively, there is clearly still some overlap. A particularly evident contradiction within current and previous findings is the role of self-regulation, self-monitoring, inhibition, shifting, and the role of the Stroop task, and why to date, between three different prominent 'best' statistical methods (PCA, CFA, congeneric) there is still no consistency.

It was suggested a reason why the CF construct explained the second amount of shared variance was because it is able to assess some skills in isolation, despite the underlying mechanisms of the construct requiring dual tasks. Dual tasks by nature require two relatively simple tasks to be carried out in close succession or concurrently. Thus, it could be argued that dual task is the interface between SAC and CF which is why there will always be significant overlap, particularly with the skills of inhibition and shifting, consistent with attention being fundamental to all cognitive performance (Sarter et al., 2001) and the correlations from the variable selection and reduction phase. This premise is also consistent with Kane, Hambrick, and Conway (2005), who suggested that not all WMC tests are dual tasks, and instead, share processes with non-dual task measures. This might explain why some tasks not considered dual in nature were related to dual task measures and vice versa.

For example, as task demands increase and require one to carry out two tasks simultaneously or in close succession thereby becoming more complex, it is these fundamental skills that begin to work in an integrative manner that evolve to become more executive in nature. Therefore, inhibition would arguably belong under CF, consistent with the differences in developmental trajectories between skills. This is consistent with Garon, Bryson, and Smith (2008) who highlighted differences between simple and complex inhibition tasks are dependent on whether or not other skills are required, such as WM. This premise is further consistent with previous definitions regarding cognitive flexibility that suggest incorporating multiple sources of information simultaneously are essential for this construct (Anderson, 1998). This might explain current findings where the placement of inhibition is contentious, and why the Stroop test was an unreliable measure under CF, because there seems to be controversy if this is related to shifting or inhibition. It could be that inhibition may be functionally separable (Dajani & Uddin, 2015; Miyake et al., 2000).

For example, if a healthy adult is required to engage in dual tasks, inhibition is a key component to Cognitive Flexibility because one must not only inhibit distractors (arguably a simple skill for healthy adults if used in isolation (e.g., SAC) but also draw on divided attention which requires one to consciously pay attention and distribute cognitive resources between two tasks at the same time (Goldstein, 2008). Thus, an increase in task complexity implicates the cognitive resources and capacity available to divide one’s attention. These skills are then used to incorporate WM to hold and manipulate information temporarily, where WMC uses controlled attention to maintain or supress WM (Engle, Kane, & Tuholski, 1999a; Engle, Tuholski, Laughlin, & Conway, 1999b), consistent with the placement of Elevator Counting with Reversal in the current study. Controlled attention is useful for any tasks that require goal maintenance, decision making, error monitoring, conflict resolution, effortful memory search or suppression of distracting information that is generally domain free (Hunter & Sparrow, 2012; as cited in Dajani & Uddin, 2015; Miyake et al., 2000), where greater WM demands are imposed when task shifting paradigms are used (Dajani & Uddin, 2015). Thus, it is reasonable to suggest that for a child, the Stroop would be considered complex and higher order, which explains why Anderson placed this under the CF construct. By contrast, the Stroop test would clearly be facile for an unimpaired adult, as evidenced by the extreme lack of variability demonstrated in the current findings (*M*= 1.68, *SD*= 0.43), and therefore would be classified as a simple task measuring controlled, intended suppression of prepotent responses (given the automatic comprehension of the words). It could be argued that the healthy sample used (which inevitably dictates a low task load) did not make many self-corrections because they successfully performed the inhibition task, thus making the complexity of the tasks questionable.

Another reason to possibly explain why Anderson's descriptions between AC and CF overlapped besides upwards extrapolation of skills is the purported differences between shifting. Miyake and colleagues' (2000) "Shifting" factor refers to shifting between tasks, mental sets, or operations, also referred to as 'attention shifting'. This involves "the ability to perform a new operation in the face of proactive interference or negative priming" (Miyake et al., 2000, p. 56). They referred to this as executive-oriented shifts regulated by the frontal lobes, compared to spatial shifting that is regulated by posterior regions (e.g., parietal and mid-brain) (Posner & Raichle, 1994, as cited in Miyake et al., 2000). This seems to be a term used to include a variety of shifting abilities, where as Dajani and Uddin (2015) discuss differences between task and set shifting and their relative placement within a hierarchy. Set shifting has been described as the ability to shift attentional control *within* a task or between features of the same stimuli, typically considered a lower level form of cognitive flexibility (Dajani & Uddin, 2015). This is consistent with the TMT-B in the current findings where one must switch between letters and numbers. Task switching, on the other hand, is more complex in nature because of the switching *between* tasks, with two different instructions (Dajani & Uddin, 2015). This was demonstrated by the Telephone Search While Counting task of the TEA. One must listen to the tones, and also selectively attend to a visual stimulus simultaneously. Thus, whilst these two skills are posteriorly mediated and classified as simple in isolation, they become anteriorly mediated and more complex because of dual nature of the task (arguably). Together, these findings may explain why previous factor analytic studies have found a mixture of findings.

In summary, the originally proposed Attentional Control construct is best defined and reconceptualised as Simple Attentional Capacity because these tasks are too simple, likely owing to a posteriorly mediated bottom up approach. It is likely that the tasks are too simple because they are posteriorly mediated and do not fill up cognitive load because of the differences in developmental trajectories between children and adults. Therefore, in healthy adults where skills are established, anything that is dual in nature should be considered higher or complex attention, and this is likely to be inherent in the Cognitive Flexibility construct, and should therefore be used synonymously with Cognitive Flexibility. Therefore, the CF construct is likely complexity driven and is measuring shifting and divided attention, inhibition (when used in dual tasks), and WM, and is underpinned by simple attentional capacity processes. Therefore, a model that is considered EF must reconcile the placement of attention. Arguably complex attention tasks could be considered EF, reserving posterior attention for different theories.

9.1.4 Information Processing

Within the Information Processing (IP) construct Anderson suggests that one must be efficient and fluent, where efficiency implicates how quickly and accurately one can produce an appropriate response. Thus, tasks entered into IP included Verbal Fluency (FAS and Animals), d2CONC, 5-point, and Rule Shift. This was a well-supported construct because all tests demonstrated consistency with respect to their loadings, however demonstrated lower than desired reliabilities. Overall this construct explained the third highest amount of variance, thereby supporting the first hypothesis.

Specifically, study findings indicated the test of d2CONC demonstrated the strongest loading. This could be because this is a combined measure of both speed and accuracy that may be able to more accurately tap into the purported skills of concentration, essentially reflecting how quickly one can take in and process information. Traditionally, the test of d2 is a measure of selective and scanning attention with elements of processing speed, and accuracy (Bates & Lemay, 2004; Brickencamp, 1981). However, it may be sensible to suggest that whilst this may represent a relatively simple visual search task akin to TMT-A or Map Search given the number of correlations across a variety of domains, it may also be reasonable to suggest that the combination of both speed and accuracy provides more value specific to this task, and therefore it rightly belongs under the IP domain. This is similar to Bates and Lemay (2004) who found their d2 concentration score represented an overall performance measure that is superior to total correct, because of both speed and accuracy measures used. It could be that despite the healthy sample making few errors on a range of other EF tasks, the number of commission errors was greater in the d2 task because this relied on quick responses under time constraints, whereas other tests were not constrained by time limits, and instead were simply timed. Essentially, the fact that participants were under time limits elicited more errors, compared to when time was simply observed or not imposed, leaving room for participants who were so inclined to be more cautious and make less errors.

Interestingly, the 5-point test demonstrated the second strongest loading. It could be that both the d2 and 5-point are non-verbal tests which seem to be the strongest within the construct. Traditionally, the 5-point is described as a measure of non-verbal fluency and executive control (Regard et al., 1982), and the current findings demonstrated that the correlations to other tests also represented elements of visual attention and shifting, and processing speed (e.g., FAS-15, d2 measures, AM time, VE).

Study findings indicated that Rule Shift was the next strongest indicator. Although typically a measure of shifting (Wilson et al., 1996), this test was placed under the IP domain because it was thought to include the ability to respond to and process information as quickly as possible. This was also placed here to meet the minimum requirements of SEM. Direct comparisons are often difficult at a subtest level because this test forms part of test battery for clinical populations and is rarely used in isolation. Testa et al. (2012) did report that she set out to use this test, however it did not make it into the final factor solution. It could only be assumed that this test was too simple for their healthy adult sample given that traditional scoring relies on errors, whereas in contrast the current study used time to complete as a measure.

Finally, study findings indicated that both fluency measures (letters and words) were the weakest and most unreliable measures within the construct. They were however correlated to each other which is consistent with previous research suggesting they load on the same factor (Hedden & Yoon, 2006; Testa et al., 2012; Unsworth et al., 2011; Whiteside et al., 2016). It is likely that these tests were most unreliable because they have previously been purported to measure a variety of skills. For example, Verbal Fluency has been proposed as a measure of attentional control because of the taxing resources in retrieval of words and mental speed (Boone et al., 1998). Furthermore, self-monitoring to avoid errors, suppression of previous responses and generating cues to access new items have also been proposed to be measured by Verbal Fluency tasks (Azuma, 2004), and to tap into WMC (Rosen & Engle, 1997). Despite the strategy scores such as clustering and switching not making it to the analysis stage of the study, it is clear that there is some element of strategy to effective performance on fluency measures (Troyer et al., 1997). This array of skills has been further supported by different factor structures across various studies. For example, a Language factor has been proposed (Whiteside et al., 2016), a Strategy Generation and Regulation factor (Testa et al., 2012), and Verbal Abilities factor (Pires et al., 2018), which have been related to other cognitive constructs such as WMC and Vocabulary (Unsworth et al., 2011). Thus, a variety of different skills may contribute to successful performance on the VF measures.

As a construct, study findings are consistent with previous research (Kail & Miller, 2006) that found processing speed is a quantifiable domain capable of measurement in its own right. For example, a Speeded Processing factor was found in Boone et al.'s (1998) factor analysis study which comprised time to complete all parts (A, B, C) of the Stroop, Verbal Fluency (FAS), and Digit Symbol in a sample (*N*=250) of mixed populations. Furthermore, Pineda and Merchan (2003) also found a Speed of Inhibitory factor comprising time to complete all Stroop measures, and a Visual Motor Speed factor comprising time to complete TMT-A and B in a healthy sample (*N*=100). Study findings are also similar to Chiaravalloti and colleagues (2003) who also found in their PCA of mixed participants (*N*=92) a separable Information Processing domain. However, they also found that different measures of information processing loaded across various factors, suggesting IP may not be a unitary construct and identified different factors relating to simple and complex processing speed. Study findings are also consistent with previous research that used latent variable approaches. For example, results are consistent with Pires et al. (2018) who found their CFA demonstrated a three-factor model comprising EF, Verbal Abilities (VA) and Processing Speed (PS) in a sample (*N*=90) of healthy adults aged 18-33. They suggested Processing Speed loads separately and indicates that EF and Processing Speed are related yet separable (Pires et al., 2018). Study findings were also consistent with Conway et al. (2002) who conducted a latent variable analysis on 120 young adults and found support for their construct Processing Speed. This construct comprised the Digit Symbol Substitution test (RT as the outcome measure), Digit and Letter Copying (number of letters copied correctly in 30s as the outcome measure) and lastly, Pattern and Letter Comparison (increasing in difficulty; total number of patterns and letters compared after 1 minute used as outcome measure). Furthermore, in their CFA, McAuley and White's (2011) study also found support for their Processing Speed factor in a large sample (*N*=147) of healthy 6-24-year-old's. Their construct comprised simple RT which was used as a measure of processing speed and RT, the Stimulus-Response Compatibility Task, which was used as a measure of processing speed and inhibition, and lastly, the Go/No-Go task which was used as a measure of processing speed and inhibition.

The IP construct has gone relatively underrated within the cognitive models of EF, not only because of the lack of internal consistency of the construct itself, but mainly because it has received little attention to be included in a model of EF at all. A review of the literature suggests that IP is rarely considered an EF and is more or less a skill that we know 'exists'. Thus, a review of the literature clearly supports the notion of 'Processing Speed' or 'Information Processing Speed' as a factor or construct, found using a variety of statistical methods (e.g., factor analysis, CFA, and in the current study; congeneric modelling). However, a possible explanation why many of the tests in the current study were unreliable could be because speed is fundamental to *all* cognitive performance, consistent with the multifactorial nature of EF. All of the tests within this construct have been purported at some stage as a measure of speed within the literature, most likely because “processing speed has been used to refer to a variety of measures that get used in different ways that may tap underlying components to varying degrees” (Cepeda et al., 2013 p. 270). Whilst it is not a novel finding that speed is fundamental to all cognitive performance because many tasks inherently require elements of speed, it is however, likely the reason why studies have proposed processing speed as a globalised function (Fry & Hale, 1996; Hale & Jansen, 1994; Peter et al., 2011), that represents a systemic mechanism that is not specific to a particular task (Kail & Miller, 2006). It is further likely the reason why some have suggested it is important to remove the influence of PS to understand the contributions of other higher order cognitive processes (Cepeda, et al., 2013), so much so, that most PS measures are simple tasks, so that the contribution of other higher cognitive functions is minimised (Fry & Hale, 2000). Thus, the ambiguity in the construct leads researchers to question what the role of PS is especially to EF, and if this is domain specific or a more globalised function. Given the lower than desired reliabilities found in this construct, it may be sensible to suggest that whilst IP was found to be a measurable construct, it may also in fact be a globalised function that recruits a variety of different skills.

It is the contention of this paper to suggest that IP, whilst not traditionally conceptualised as an Executive Function, is a critical foundation skill of overall EF. In this study and more broadly, IP is a construct that is consistently measured using EF type paradigms (e.g., Cepeda et al., 2013; Chiaravalloti et al., 2003). Therefore, this measurable construct extends beyond simple RT measures as indicated in some previous studies (e.g., Conway et al., 2002; McAuley & White, 2011), offering more than baseline measures. It could be argued that the IP construct shares similarities between SAC and therefore is best conceptualised essentially as a ‘simple’ construct, however these constructs should not be considered similar beyond the notional level of complexity. The study findings argue strongly, that whilst attention can be demarcated based on complexity, where more simple attention should not be considered 'executive' and more complex attention is arguably better defined as 'Cognitive Flexibility', IP when measured beyond simple RT, is unequivocally an Executive Function. Furthermore, it is likely an EF that permeates all aspects of cognition, hence the generalised nature of its definition in previous studies. In any model of EF, especially the Anderson's Executive Function model, it is certainly evident that IP provides additional explanatory value, and strong validity, as all tasks under IP were strong indicators and none were removed.

In summary, whilst this study has found that we may be able to distinguish between healthy people on the basis of their performance relative to speed, it is the value of the IP domain that was a well-supported construct that did not rely on observations of time taken, and instead, relied on time constraints. This was an interesting finding because all other measures across a variety of domains in the current study that had an element of speed were not imposed by time restrictions (except for Map Search under SAC). However, all variables (besides Rule Shift) underpinning the IP construct measured how much a person could complete under time restrictions. Therefore, the IP construct which focussed on the production of output under time constraints, and where the level of complexity of a task dictates the speed with which it could be completed seems to be of value.

Therefore, these results suggest that the IP construct is capacity over complexity driven, because this domain can be evaluated by the speed (how quick) with which an individual is able to process, respond to task demands, and produce information (both quality and quantity) in a strategic manner. Arguably this becomes about 'how much' one can process under time constraints efficiently and fluently, which this relies on intact frontal systems (Anderson, 2002). Essentially, efficiency within this construct refers to faster performance with more correct responses, where a lower score indicates better efficiency.

Reaction time measures are not useful in a model of EF, because RT is typically referred to as the minimum amount of time required to produce a correct response (Pachella, 1974). Instead, the IP domain is best measured using traditional EF tasks, emphasising the opportunity to produce output under time constraints. Thus, consideration of Information Processing as a construct that represents efficiency, fluency and speed of processing is considered a strong and unique finding.

9.1.5 Goal Setting

Goal Setting as a construct reflects initiative, conceptual reasoning, planning, and strategic organisation. Tasks entered into Goal Setting (GS) were as close to the theoretical placement recognised by Anderson, however only three tests were exactly the same. Tasks entered into the model were Zoo Map, Key Search, ROCFT, AM, TOH, and PA. Not all purported skills (e.g., initiation) were captured by the included tests and many tests demonstrated very weak loadings and reliabilities, which is why this construct explained the least amount variance overall. Reconciling the mathematical limitations of including the set of tests in this construct meant that poor indicators were retained. However, only one test met the minimum technical requirement for SEM however if all were removed, the construct itself would not be validated. Whilst poor fitting measures were evident Goal Setting demonstrated a significant model fit, and thus the latent construct was confirmed by congeneric modelling.

For a researcher and clinician alike, the inability to tease out individual skills is a contentious issue. For a researcher, understanding each skill on an individual basis would be beneficial because skills could be mapped out more precisely, which would therefore inform clinical practice. Whilst this may be achieved for lower level skills, task impurity implicates skills considered higher order.

Thus, a possible reason why Goal Setting was a challenging construct to define is because Goal Setting is a construct that is intangible. A review of the literature highlights that of the models reviewed, goal setting/formation is always the outcome and never a concept, and there is a paucity of validation of this construct. This might be because in order to assess the ability to formulate goals, it is the accumulation of skills lower in the chain of processing, effectively representing a hierarchy of skills. For example, to set a goal, one must plan, and in order to plan one must use working memory to temporarily hold and manipulate information, inhibit competing or distracting stimuli, and divide and shift attention. Thus, skills preceding the GS construct such as those in CF, AC, and IP could be thought to play a role in effective execution of a goal. The only model which provides support is the problem-solving framework (Zelazo et al., 1997) which highlights the hierarchical steps necessary to formulate and execute a plan, thereby the only model that conceptualises Goal Setting. However, there is a failure to consider which of these skills at the bottom of the hierarchy are necessary and how they work within a set of skills to contribute to EF as a ‘macro construct’. For example, a person may be capable in completing complex tasks, however only when asked to do so, and simply lack initiative where it does not occur for them to do anything (Lezak, 1982). Therefore, previous approaches to validate this construct have failed because mathematical procedures will never support a higher order task or construct due to task impurity, and it is therefore essential to use theory as a guide to group skills to represent this construct.

Specifically, study findings indicated the Austin Maze was the strongest indicator and demonstrated the strongest reliability, suggesting the AM to be a very good fit within Goal Setting. The AM is a contentious choice for test selection given it is often considered a learning and memory task. That said, the nature of the task elicits planning and regulation by manner of errors and self-corrections. Thus, a ‘plan’ is not necessary to complete the task, but efficient skills necessary in overall complex and novel behaviour is essential. For example, Testa and colleagues (2012) found the Porteus Maze errors loaded onto a Prospective WM factor with the TOL-R. Furthermore, Pietrzak et al. (2008), found a Learning Efficiency factor using the Groton Maze Learning Task total time and efficiency index, as well as an Error Monitoring factor when the use of perseverative, legal, and rule break errors were used. Errors are central to performance on the maze task, and it might be that time was used in the current study and not errors like previous studies causing conflicting results.

Study findings indicated the TOH was an arguably strong indicator in the congeneric model, however the reliability index was poor. Much like other measures included in this construct, there are a variety of skills that have been purported (e.g., planning, inhibition, WM, spatial WM, information processing speed) for the TOH. Previous research reporting CFA results reported the TOH total number of moves loaded strongly to an inhibition construct (Miyake et al., 2000). Furthermore, EFA results demonstrated total time and total correct loaded onto a prospective WM factor (Testa et al., 2012). The outcome score is likely to be dictating the difference in study outcomes, as 'time' was the only outcome measure used in the current study, where previous studies utilised number of errors or moves. It could be that time is confounded by other variables reducing the reliability of this measure. Given the strong loading of the test, and strong theoretical support for its inclusion, it is reasonable that the TOH arguably fits as a measure of planning within the Goal Setting construct, as evidenced by previous literature with different versions (TOL-F; Debelak et al., 2016), or childhood populations (Levin et al., 1991). Furthermore, Welsh et al. (1995) did not find planning significantly contributed to performance on the TOH in healthy adults quantitatively, they did however *qualitatively* report that planning and mentally evaluating requirements of tasks were used in order to reach sub goals, which necessarily increased time to avoid errors. Thus, whilst time may be confounded by other variables, it might also be a more useful measure that captures a variety of skills inherent in a task, as opposed to just total moves that does not consider the time to complete. It is evident that a variety of skills is necessary for effective performance on the TOH.

Study findings indicated the ROCFT was a weak and unreliable indicator. Whilst the derived score was used, the variable selection and reduction phase performed initially indicated that traditional scores elicited very few correlations with other measures and therefore the value of the ROCFT is likely obscured by inefficient scoring procedures. Traditionally, the ROCFT has been used extensively in both research and clinical practice in paediatric and clinical populations in Australia, however, a search of the literature identified scarce factor analytic data relating the ROCFT, in particular with healthy adult samples. Meyers and Meyers (1995b, as reported in Strauss et al., 2006) conducted a PCA in healthy individuals and found a five-factor solution. However, the only variable similar to the current study was the time to complete, which loaded onto a Processing Speed factor. Thus, caution is advised when using this as a measure of strategic organisation in healthy adults.

Similarly, Key Search and Zoo Map demonstrated the same pattern of findings. Traditionally these tests form part of the BADS test battery which has been used extensively in clinical practice, and individual subtests are rarely used in either research or clinical practice. Bennett et al. (2005) explored the BADS using a clinical population and found the Zoo Map test loaded onto a factor with other measures and likely reflected sequencing and self-monitoring. They also found that Key Search loaded onto a factor with other measures labelled as initiation and cessation and control of action. However, interestingly no planning or goal setting factors were found when these are presumed to be tests of planning. The one healthy sample that used these measures was Testa et al. (2012). However, their Zoo Map test loaded onto a factor in which they labelled as Task Analysis along with the Brixton test, CNT, WCST (FA), and Cognitive Estimates.

Lastly, study findings demonstrated that PA was also a weak and unreliable measure.

The error variance was however correlated with the TOH, suggesting there may be an element of shifting of sequences or planning common to both.

With respect to skill assessment, it was originally anticipated that the alternate scoring of the verbal and non-verbal fluency measures would clarify the subskill ‘initiation’. However, traditional scoring provided more value, and therefore there were no tests available to test this subskill. Theoretically however, this skill does belong under the Goal Setting construct (Lezak, 1982).

In summary, this thesis has suggested that as a construct increases in complexity, the ability to define and assess it becomes more difficult and will remain difficult to operationalise which might be why there is a paucity of validation of this construct within the literature. This is because task impurity implicates the ability to break down the multifactorial skills required to work in an integrative way to reach the end state goal. Whilst other skills that are defined by dual tasks such as those in the CF construct also have this problem, it is evident that this problem persists when tasks increase in complexity that require the integration of even more than two skills simultaneously. Thus, Goal Setting is arguably the most complex construct, because within this are the proposed skills of planning, strategic organisation, initiative and conceptual reasoning, and are thus termed higher order despite not being able to effectively assess initiative. Therefore, these results are unique because a formidable construct that has remained challenging to define and validate has been offered and supported statistically with a strong theoretical basis. Goal Setting as it is operationalised in this study, reflects the highest level of complex EF able to be assessed by existing measures. In this way, Goal Setting should be considered at the apex of complexity, with other latent constructs beneath it conceptually.

9.2 Assessing the Predictive Validity and Interrelationships of Verified Constructs in a Model of EF in a Healthy Adult Population Using Regression Analyses

In order to address the overarching aims of the study to bring clarity towards EF, a series of regression analyses were run to identify the strength and direction of the relationships between constructs. Study findings indicated all constructs were unique in their contribution and played a significant role within a model of EF, thus providing support for the fractionation of EF and shifts the once applied premise of viewing EF as a homunculus.

Path analysis using SEM would have provided more robust conclusions regarding causality amongst the set of variables, though given limitations imposed by sample size and constitution, regression analyses were conducted instead. Study findings indicated that SAC only contributed statistically to CF. Therefore, the second hypothesis that stated Attentional Control would be the strongest predictor, explaining the greatest variance in all other latent constructs was not supported. It was the contention of this paper to argue the role of attention in a model of EF. Attention is a skill that is fundamental to all cognitive performance, and a substantial body of literature has demonstrated the importance of considering this as a construct in the overall conceptualisation of EF, particularly the way in which it manifests itself in a range of disorders with an inherent executive dysfunction component. Thus, it was one of the many reasons why Anderson's model was selected to be validated in healthy adults. Needless to say, this thesis has demonstrated that the reconceptualised construct of Simple Attentional Capacity was removed from the model of EF with good theoretical justification. It is likely to have elicited a bottom up approach that ultimately should not be considered exclusively as an Executive Function and is likely owing to the fact that developmental trajectories differ between children and adults. However, despite the use of adult level tests it was evident that this construct still represented Simple Attentional Capacity. Therefore, in healthy adults 'complex attention' or 'attentional' control is likely a core component underpinning the CF construct and was difficult to separate because these skills are established and not emerging like they are in paediatric samples.

Study findings indicated that as a construct, Cognitive Flexibility explained the most amount of variance. It was also interesting that the unique contribution as indicated by partial coefficients demonstrated that IP and CF held the strongest relationship to each other. Therefore, it could be argued that these two are considered important constructs in the overall conceptualisation of EF. Furthermore, given the significant positive relationship between CF to GS, it could also be implied that a hierarchy of skills is evident in the overall conceptualisation of EF. For example, this positive relationship suggests that as one's ability to shift and divide attention, inhibit irrelevant stimuli in order to use WM resources to maintain and manipulate information, so too does their ability to set goals and plan out how to achieve them effectively in an organised fashion. Essentially, those tasks that are dual in nature, should facilitate effective goal formation. Thus, those skills below in the chain are drawn upon for efficient goal directed behaviour, which therefore strengthens the current study's proposal that skills underpinned by CF are important factors contributing to effective performance for GS.

To that end, it appeared that Information Processing was the strongest and therefore most influential predictor across all constructs (CF and GS). This provided partial support for the third hypothesis, because Information Processing was not the second, but rather, was the strongest predictor of other latent constructs, therefore demonstrating that IP is an influential component in a model of EF. Furthermore, a negative relationship was found between both IP and CF, and IP and GS which is consistent with Anderson's premise that bidirectional relationships may account for the enhancement or decrement in tasks that mediate performance at a higher level. For example, slowed reaction time and delayed responses are impairments of this construct which may or may or may not be accounted for by poor organisation which is fed to the Information Processing sub domain. A possible reason why IP was the strongest construct could be because this is a capacity bound construct, whereas CF and GS are complexity bound constructs. The more efficient one is at processing, responding to task demands, and producing output strategically under time constraints, should facilitate their ability to effectively carry out dual tasks, and set goals in an organised manner.

This thesis has demonstrated the importance of Information Processing. Rather than attention being a central component towards a model of EF, IP was found to play a significant role in its contributions to other higher order constructs. Therefore, this thesis argues that Information Processing should be considered a foundational skill lower on the continuum of performance, because it holds a strong affinity to higher order Executive Functions.

Study findings demonstrated that GS as represented by the current study was the only construct that was the most ‘latent’ and ‘intangible’, and hence difficult to assess, and it may be why theorists have had difficulty operationalising the construct. However, this thesis has found that skills below in the chain of processing, specifically those under the Cognitive Flexibility and Information Processing constructs are influential skills that contribute towards effective Goal Setting. Thus, in terms of complexity, if the skills underpinned by Cognitive Flexibility were not considered important when assessing Goal Setting, then the relationship between CF and GS would be of a lower strength. The parallel argument therefore, is that this overlap would simply imply that GS is in fact the same construct as CF, in which case these constructs would demonstrate a stronger relationship to the point of multicollinearity. This was not the case in the current study. In contrast, in terms of capacity, the same premise could be applied for IP and its influence to GS, where this also did not demonstrate a weak relationship, nor did it demonstrate multicollinearity. Therefore, these results confirm that

EFs are both unified and diverse, consistent with Miyake et al.'s (2000) findings and move away from the once applied premise of viewing EF as a homunculus.

In summary, with all of the caveats outlined earlier in mind, this gave rise to the mediation analysis, where the removal of SAC was warranted and the relationships re-investigated. These results confirm that the simplicity of Attentional Control as represented by more posterior attention tasks, does not significantly explain EF performance at a higher level. SAC is redundant in a model of EF in healthy adults and complex attention is required instead, where complex attention may be represented by CF (only one of the many necessary constituent skills required for CF). Furthermore, considering GS has been identified as the epitome of Executive Functioning and at the top of the hierarchy, it seems reasonable to apply this construct as the outcome or end state goal consistent with most definitions of EF, where complexity bound skills lower in the chain of processing predict performance at a higher level (e.g., CF). Lastly, considering the capacity bound IP construct was the strongest predictor in all models, this warranted the inclusion of the speed at which one is able to process and respond to task demands and produce output strategically as a mediating factor in performance (e.g., GS).

9.3 Investigating Information Processing as a Mediating Factor in Cognitive Performance

The indication that Information Processing shared significant associations and predictive validity of both higher order constructs warranted further investigation. The theoretical evidence also supported the notion that IP may in fact be a mediator of higher order performance, much in the way that Attentional Control was also expected to be critical in the overall framework. Justification for the omission of AC for consideration as a mediator has been addressed, though to reiterate, AC as conceptualised in the current study, is likely to reflect a very simple, bottom-up process not essentially or exclusive to EF, but rather essential to cognition more generally. This thesis has highlighted the importance of considering speed, accuracy, and overall efficiency as a means of understanding higher order executive performance, and mediation analysis facilitated testing this theoretical supposition.

Mediation specified both direct and indirect pathways. Thus, the aim of the study was to explore the relationship between Cognitive Flexibility and Goal Setting to see if this was mediated by Information Processing. Results indicated that IP partially mediated the relationship between CF and GS. On its own, increases in CF led to higher GS, however when controlling for IP, the strength of the relationship was reduced between CF and GS consistent with partial mediation.

9.3.1 Direct effects

Mediation analysis yields three direct effects, CF and GS (*path c*), CF and IP (*path a*), IP and GS (*path b*), and an overall indirect effect with IP mediating the relationship between CF and GS (*path c'*).

With respect to path c, Cognitive Flexibility demonstrated a significant positive relationship with Goal Setting. This means that as one's ability to shift and divide attention, inhibit irrelevant stimuli in order to use WM resources to maintain and manipulate information, so too does their ability to set goals and plan out how to achieve them effectively in an organised fashion. Essentially, those that are able to successfully complete tasks dual in nature, are better able to execute goal directed behaviour. This therefore provides support for the skills underpinned by CF as important factors contributing to effective performance for GS.

Study findings are consistent with previous theories that conceptual inflexibility results in the inability to shift, disengage attention and perseverate, in turn affecting rigidity in the ability to solve problems and therefore achieve goals efficiently (Lezak et al., 2004; Pennington & Ozonoff, 1996). Similarly, this is also consistent with findings suggesting individuals struggle with new procedures and fail to adapt to new demands if the Cognitive

Flexibility construct is not intact (Anderson, 2001), all of which are necessary when examining Goal Setting. Moreover, a variety of research findings relating to autism also support this relationship. Those with autism typically display impairments relating to cognitive flexibility (Yerys et al., 2009). That is, they demonstrate the inability to shift their attention (Gioia, Isquith, Kenworthy, & Barton, 2002) and rigidity is evident particularly when difficulties arise when plans must be adapted in everyday life (de Vries & Geurts, 2012). Importantly, study findings are consistent with Lezak's conceptual model of EF, particularly the organisation of the distinct categories necessary to formulate a goal (capacities necessary for formulating goals, planning, carrying out the plans to reach the goals, and performing these activities effectively). Lezak highlights that in order to carry out the activities it requires the capacities to "initiate, maintain, switch, and stop sequences of complex behaviour in an orderly and integrated manner" (Lezak, 1982, p. 290). One must also "conceive alternatives, weigh and make choices, and entertain both sequential and hierarchical ideas necessary for the development of a conceptual framework to the carrying out of a plan" (Lezak et al., 2004 p. 614). She also highlights that impulse control, memory and sustained attention are also necessary for the completion of a goals, where failure to complete a goal is because a failure of one or more of the capacities required to complete a goal are impaired (Lezak et al., 2004). Furthermore, the fact that working memory, inhibition, and cognitive flexibility have been classified as core EFs, and reasoning, problem solving and planning have been classified as higher level EFs (Diamond, 2013) where higher level EFs require the maintenance and manipulation of information to organise goal directed behaviour in a strategic manner (Vestberg, Reinebo, Maurex, Ingvar, & Petrovic, 2017), this further strengthens current findings of Goal Setting as being at the top of this hierarchy. This is also consistent with those that suggest EFs at the highest level require executive working memory, because this is required to simultaneously recruit, coordinate and direct behaviour towards future goals (Luciana et al., 2005). The value of these findings is that it further strengthens the proposition that skills below in the chain of processing are required for efficient execution of a goal. Therefore, these findings support the notion that there is a hierarchy of skills in the chain of overall processing consistent with the skills, and directionality outlined in Anderson's model, thus confirming the nature of the relationship of Anderson's model of Executive Function in healthy adults.

With respect to path a, Cognitive Flexibility demonstrated an inverse relationship to Information Processing. This means that as one's ability to shift and divide attention, inhibit irrelevant stimuli in order to use WM resources to maintain and manipulate information, the speed with which an individual is able to process, respond to task demands, and produce output in a strategic manner efficiently becomes faster. Essentially, greater Cognitive Flexibility facilitates the rate of Information Processing.

Previous researchers have concentrated their efforts to investigate the links between processing speed, WMC (as part of CF), and general fluid intelligence, suggesting there is a bidirectional relationship, or at least substantial overlap (Ackerman, Beir, & Boyle, 2002; Kyllonen & Christal, 1990). However, the directionality between processing speed and cognitive abilities remains elusive in the literature, specifically, whether WMC sets limits on processing speed, or vice versa where processing speed limits WMC. For example, some have suggested the most important resource towards efficient WM and general intelligence is information processing speed (Fry & Hale, 2000; Kail & Salthouse, 1994, as cited in McAuley & White, 2011), perhaps because a faster processing time means there is less time to have to hold information in WM, or because speed determines WMC because it takes time to process (e.g., encode, transform, retrieve) (Conway et al., 2002). Conversely, others suggest WMC predicts processing speed (SüB, Oberauer, Wittmann, Wilhelm, & Schulze, 2002) perhaps suggesting WMC limits the ability to process information quickly (Ackerman et al., 2002).

In one study it was suggested that those who demonstrate a higher speed of information processing in higher order cognitive processes were able to do so because they demonstrated superior cognitive abilities, in particular those underpinned by CF such as WM, inhibition, and shifting. For example, although the use of diffusion models in their study, and measuring general intelligence and not EF, Schubert, Hagemann, and Frischkorn (2017) found “more intelligent individuals process information faster specifically because of faster higher-order processing” (p.1506). The authors suggested that this was because participants were able to inhibit extraneous processes quicker, in turn “facilitating the transmission from frontal attention and Working Memory processes to temporal-parietal processes of memory storage (Polich, 2007, as cited in Schubert et al., 2017, p.1506). It could be suggested that WMC is a predictor of processing speed because speeded tasks might place demands on WMC (Conway et al., 2002). To this end, Conway et al. (2002) used simple processing speed measures and found that this was not related to WMC. This therefore strengthens current findings that IP must be considered on a higher order basis to be relevant in a model of EF, and it could be because it does place demands on WMC, divided and shifting of attention akin to the CF construct, which might be why the Verbal Fluency measures (FAS, animals) that rely on attentional control because of the taxing resources in storage and retrieval of words (Rosen & Engle, 1997; Troyer et al., 1997) were unreliable in the congeneric model, and might be why there was such a strong relationship between IP and CF.

Taken together, study findings are consistent with previous literature that suggest a bidirectional relationship is evident between CF and IP, where WMC is underpinned by the CF construct. Given the findings of the mediation analysis according to the strength of direction of CF predicting IP (b= -7.18), study findings could be implied to suggest that Cognitive Flexibility where WMC is subsumed, does in fact set limits on Information

Processing speed abilities because of the demands on WMC within the IP construct.

Lastly, with respect to path b, Information Processing demonstrated an inverse relationship to Goal Setting. This means that poor performance on the speed with which an individual is able to process, respond to task demands, and produce output in a strategic manner efficiently, their ability to set goals and plan out how to achieve them effectively in an organised fashion decreases. Essentially, efficient processing of information facilitates goal directed behaviour.

Evaluating these abilities under time conditions is clearly valuable, because it provides another lens to investigate the influence of IP to higher order constructs. This premise is supported through a study that found planning deficiencies in MS patients were attributed to information processing speed, only when time restrictions were imposed (Owens, Denney, & Lynch, 2013). The authors concluded that when sufficient time was allocated and no time restrictions were imposed, MS patents were able to successfully complete the TOL, however when time restricted, performance was affected.

Thus, given the relationship between CF and IP has been demonstrated, it seems sensible to imply that the relationship between the two both facilitate goal directed behaviour, thus confirming the integration of cognitive constructs required for efficient Executive Functioning. This ultimately leads to information processing speed as being a predictor of EF, because processing speed tasks require EF (Diamond, 2002). It could be suggested that IP is a construct incorporating the speed with which one can process, respond to, and produce output in a strategic manner that efficiently facilitates effective goal directed behaviour, possibly because of the demands of strategy use and WMC subsumed by the IP construct that are clearly required to execute a goal, and Goal Setting clearly draws on these skills.

9.3.2 Indirect effect

Notwithstanding the direct effects indicated, in the presence of a significant model of mediation, the interaction between predictors is of greater importance and overall more relevant. With respect to path c', results indicated that IP partially mediated the relationship between CF and GS. On its own, increases in CF led to higher GS, however controlling for IP, the strength of the relationship reduced between CF and GS consistent with partial mediation. These findings highlight the importance of efficiency of Information Processing and its influence towards other skills within a model of Executive Functioning.

Whilst there are differences between the measures used across studies, current findings are relatively consistent with those that found speed as a mediating factor in many cognitive tasks. In particular, Processing Speed has been identified to be an influential factor in age-related decline in cognitive performance, and other cognitive processes. For example, Bunce and Macready (2005) found their Processing Speed construct as represented by Digit Symbol Substitution test and Choice RT variables via PCA, accounted for age difference in remembering and knowing, but executive control measures had little influence in both young and old adults. This suggests that speed has a larger influence than does executive control in remembering a knowing. Similarly, after controlling for PS as a composite measure, Lövdén, Rönnlund, and Göran-Lars (2002) found age-related differences in remembering and knowing were removed, suggesting that a reduction in processing speed contributes to the age-related differences. Furthermore, Genova, DeLuca, Chiaravalloti and Wylie (2013) found performance on EF tasks in MS patients is highly dependent on PS because once PS was removed and corrected, significant group differences between clinical and control participants dissipated. Moreover, Fisk and Sharp (2004) demonstrated that accounting for Information Processing Speed, removed most of the variance accounted for by age. Because of this, age was no longer a significant predictor of their factors of Updating, Inhibition and

Access, however not for Shifting. Whilst Information Processing Speed removed half the age-related variance for Shifting, age remained significant. Lastly, Span, Ridderinkhof, and van der Molen (2004) found that differences between age groups (children vs. adults, adolescents vs. adults, and seniors vs. adults) on EF tasks (response inhibition and working memory) were no longer significant once PS was controlled for. Their findings suggest EF skills were mediated by the efficiency of a participant's ability to process information.

The current results indicate that Cognitive Flexibility and Goal Setting were related, however when Information Processing was introduced, the relationship between Cognitive Flexibility and Goal Setting reduced indicating that the speed with which one is able to process, respond to task demands, and produce output efficiently is a mediating factor in Executive Functioning performance. Therefore, IP as a mediating factor in cognitive performance is consistent with the definition of PS being an "underlying cognitive efficiency at understanding and acting upon external stimuli, which includes integrating low level perceptual, higher level cognitive, and output speed" (Shanahan et al., 2006, p. 586). Furthermore, study findings are consistent with Anderson's premise that CF and GS cannot be assessed without consideration of IP. By investigating the mediating role of IP, the present study offers a more comprehensive understanding of how IP operates within a model of EF, particularly with the associations between CF and GS. Thus, IP is a determinant for EF efficiency because IP mediates higher order cognitive processes (Pires et al., 2018), thus confirming Anderson's construction of the bidirectional relationship between IP and EF (Anderson, 2002). Therefore, IP as represented by the current findings plays a critical role in performance where speed, efficiency, and fluency may be the answer to differential EF performance.

9.4 Limitations and Future Directions

Demographic constitution of the study sample reflected a fairly homogenous group with respect to SES and intelligence, and therefore results of the current study are able to be generalised to the wider population. However, the discrepancy in number of male and females is noteworthy, where there were more females used in the study. Even though the literature does not support sex differences in cognitive performance in EF measures, a more even composition would have been beneficial. Moreover, future research could investigate cultural and individual differences in the development of a model of EF considering these variables were not considered in the present study.

Furthermore, age is a relevant variable in EF and given the large age range in the current study, this could have potentially influenced findings. In particular, given that a decline in speed as age increases has significant effects on cognitive functioning (Salthouse, 1996), and over 56% of the sample were between the ages 20-29, and 27% within 30-39, this caveat should be acknowledged. It could be why IP was the biggest influencing factor in a model of EF, because participants 'speed of processing' was relatively stable within a 20-year peak period. Future researchers should therefore examine this model utilising a sample of older adults, where cognitive decline in speed of processing is usually demonstrated.

Of course, sample size was a limitation. Although sample size was deemed appropriate for the current study, a larger sample size would be able to demonstrate greater strength of the individual loadings and therefore overall model. Sample size was a contentious issue given the large test battery taking approximately 3 hours to complete. However, one of the many strengths of the current study was the ability to take a large test battery and reduce this data to manageable form, thereby aiding test selection. Therefore, these findings will be beneficial for future researchers to be selective in their measures and narrow down test selection based on the current findings. Therefore, this study can be replicated in not only a sample of healthy adults, but also older populations using the selected tests, which would mean that the time to complete the test battery would be significantly reduced, in turn aiding the ability to obtain a larger sample size.

This thesis has demonstrated that there are multiple ways in which to measure EF, supported by a variety of tests underpinned by a variety of skills, consistent with EF theory. It is acknowledged that interpretation of current findings is constrained, particularly when the labelling of skills, tests, and constructs alike also contribute to the proliferation of taxonomies, much the same way as previous research. Thus, it is not to say that congeneric modelling is the best way, however it is a step in the right direction in obtaining a clear conceptualisation of the skills to comprise Executive Functioning.

In all variables measured for research, variability of scores is essential to uphold assumptions for statistical testing. The issue of outcome scores based on errors or time exemplifies this. Errors by nature will elicit a binary outcome (correct or incorrect, yes/no) which drastically limits variability. In contrast, time encapsulates an almost infinite spectrum of performance, and this variability is necessary for statistical testing. However, it seems that regardless of the type of outcome measure used for a higher order task, partitioning variance will always remain difficult because of the multifactorial skills subsumed for effective performance on a complex task. This was demonstrated by the correlations in the variable selection and reduction phase because a bidirectional relationship and previous studies that partition variance that is non-existent with tasks that are higher order will never suffice. This is most likely why factor analyses designed to reduce a set of tests have failed. For example, although a critique of Miyake et al.'s (2000) study was that they did not include a holistic approach to test selection, and were limited in the number of tests used which resulted in the inability to find a planning factor, Testa et al. (2012) who did use a holistic approach, also did not find a goal setting or planning factor, despite the use of tests specifically designed to measure this construct (e.g., BADS tests). It is therefore a strength of the current study that used a modified approach, with a balance of mathematical and theoretical principles, and supported the construct validity of Goal Setting despite all barriers.

Furthermore, EF theory is mired to using task impurity as an explanation for all the shortcoming of EF assessments. However, the lack of shared variance between tests is valued because one test is likely to reflect numerous skills, which is why statistically partitioning variance will always remain challenging. Rather than viewing the limited amount of variance explained by a construct as a shortfall of EF theory, perhaps a paradigm shift is necessary where the application of the pejorative term 'task impurity' can shed some light to what this actually means for researchers. We would expect the shared variance explained and reliabilities of individual variables to be lower than usual for higher order EF constructs because of the multifactorial nature of EF that require many different, yet related, EF skills. Therefore, when a construct fails to explain a large proportion of variance it should be clear that these tasks are more executively driven, and perhaps not poor fitting measures because of the array of skills required to complete a higher order task. Thus, the premise of shifting the traditional view of understanding these as poor measures is necessary because of task impurity is evident in the GS construct. Instead, the lack of variance is valued because it signifies that an array of skills are subsumed for effective performance and are thus termed higher order.

One of the main implications of the study is the arbitrary classification of the AC and CF construct. This thesis has identified that Attentional Control is best represented as a constituent of Cognitive Flexibility, most likely attributed to Simple Attention and Concentration being too simple for healthy adults since the cognitive demands did not require more complex skills, where IP was instead found to be more strongly related to CF and GS. However, one should note to not discount the importance of attention because a mediation analysis does not infer causality. This contention of the importance of attention is strongly argued throughout this thesis, where the expression of attentional mechanisms is fundamental to all cognitive performance. It is likely a result of task impurity that Attentional Control was unable to demonstrate significant strength towards a model of EF, and it could be argued AC's independence in the model reinforces its importance in models of EF. The placement and consideration of AC in the model of EF is a contentious issue and warrants further investigation particularly as it relates to the importance and role of such skills across the lifespan. Future research should therefore investigate whether the relational effects also exist in a paediatric sample, and whether AC's influence is more strongly related to CF and GS than IP given developmental trajectories. Furthermore, this model should also be validated within in a clinical sample of healthy adults to determine if the same pattern emerges given the differences between healthy and controls and the amount and types of errors made on neuropsychological tests of EF.

The current findings have significant clinical implications for the implementation of reliable and valid assessment methods to aid the diagnosis and intervention procedures of a range of adult disorders with an inherent executive dysfunction component, or dysfunction secondary to brain damage. A reliance on a variety of skills are not only necessary, but essential, for effective execution of any goal directed behaviour to help us navigate our social world. Thus, these findings provide further understanding of the complex cognitive processes associated with EF. There is no 'best' way to measure EF because multiple skills are used in different ways. However, this study can help identify which skills comprise EF and highlights the importance of the relationships between constructs that work in this integrative manner, in turn helping us understand the way in which the mind coordinates complex cognitive processes.

9.5 Conclusion

By building on the work of Anderson (2002), this study has made significant advances toward a hierarchical model of EF and more importantly, the mechanisms critical to efficient functioning at the highest level of complexity. This paper has highlighted that Information Processing is an influential construct to Executive Functioning. There has never been a doubt that processing speed is a useful and relevant factor when examining EF, there was just a lack of consensus on how it contributes to EF, or how to measure it. This often led to the conclusion that speed is just an inherent skill in many EF tasks that we 'know' we should consider. This often resulted in researchers and clinicians asking the question of whether this is a valuable construct to consider at all when assessing EF. It is with these findings that allows the complexity of IP to examined, which serves to address the contention within the literature of removing the influence of IP in order to obtain a pure measure of higher order skills, or vice versa, removing the influence of higher order functions in order to obtain a pure measure of speed (Peter et al., 2011). Whilst this need may be evident, it remains difficult because many tasks of EF inherently have a processing speed element whether timed or not, as evident by study findings. Thus, study findings conclude that time should always be used as a measure in a healthy population. Therefore, the advantage of the present findings is that the development of a model of EF in healthy adults via mediation analysis allows us to control for the influence of IP so as obtain a measure of complex higher order tasks, yet also understand the pivotal role IP plays in an overall model of EF. Thus, study findings that found three domain specific modules that are functionally specialised, yet clearly distributed and interconnected; Cognitive Flexibility, Information Processing, and Goal Setting, sufficiently explains Executive Functioning. Whilst distinct constructs have been offered, the additional value of this study is that a hierarchy of skills, working in an integrative manner in the overall chain of processing skills have been confirmed, consistent with many EF theories proposed (e.g., Barkley, Lezak, Zelazo, Anderson). Thus, although findings have demonstrated evidence for a bidirectional relationship amongst all three constructs, it is clear that they all hold their own value.

Finally, the work of Peter Anderson (2002) over a decade ago has proven an exceptional platform from which to explore definitions, tests and constructs of EF. This thesis has reconciled various issues highlighted within the literature and offers the following conclusions:

- Tests of attention, which are demonstrably mediated by more posterior cortical networks are no more exclusively associated with EF as they are other cognitive domains. As such, the constructs of attention and EF need more distinct boundaries to facilitate test selection and construct assessment. The present study findings reflect that attention tasks that are more complex as a consequence of dual processing demands, should be framed as 'Cognitive Flexibility' – demarcating them from the construct of Attention.
- Information Processing is a construct that is largely considered a general or foundation skill of cognition at large, however current study findings would encourage future researchers to consider that IP is fundamentally an EF skill. Given that the measurement of IP (beyond simple RT) is best achieved by traditional EF tasks supports this contention.
- Our ability to assess highest order EF insofar as it is operationalised as Goal Setting in the current study requires further research. A possible first step would be to conduct microanalyses (such as EFA) on each complex measure to understand and combine complex skills for a more robust measure.
- Distinguishing between those cognitive domains classified as capacity (such as IP) and those considered as complexity (e.g., CF, GS) provides a mechanistic view of *how* the inter-operationalisation of these cognitive domains work to execute goal directed

behaviour. Thus, whilst IP may serve to facilitate both CF and GS, it is implied that efficiency and fluency are best assessed at a ‘capped’ level that dictate the way in which one handles complex information.

Taken together this thesis has answered at the broadest level, what the overall nature of EF with respect to what skills it comprises, how best to assess them, and the manner in which latent constructs relate to one another to explain goal directed behaviour. Executive Functioning is best operationalised as three factors with a mechanistic view on how IP mediates goal directed behaviour.

Therefore, although these findings represent some similarities to Miyake and colleagues’ (2000) three-factor model, it is clear from study findings that a more thorough inclusion of both non-EF and EF skills provides a comprehensive analysis, and therefore conceptualisation of the nuances of skills required for efficient Executive Functioning.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*(4), 567–589. doi:10.1037/0096-3445.131.4.567

Adams, A. M., & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research*, *38*(2), 403–414. Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=7596106&site=ehost-live

Allen, P., & Bennett, K. (2012). *SPSS Statistics A Practical Guide Version 20*. South Melbourne, VIC: Cengage Learning.

Allison, P. D. (2003). Missing Data Techniques for Structural Equation Modeling. *Journal of Abnormal Psychology*, *112*(4), 545–557. doi:10.1037/0021-843X.112.4.545

Allison, P. D. (2012). *Handling Missing Data by Maximum Likelihood*. Retrieved from http://www.statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: are they separable? *Child Development*, *77*(6), 1698–1716. Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=17107455&site=ehost-live

Allport, D. A. (1980). Attention and performance. In. G Claxon (Ed.), *New directions in cognitive psychology*. London: Routledge and Kegan Paul.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: a meta-analytic review. *Neuropsychology Review*, *16*(1), 17–42. doi:10.1007/s11065-006-9002-x

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423. doi:10.1037/0033-2909.103.3.411

Anderson, P. (2002). Assessment and Development of Executive Function (EF) During Childhood. *Child Neuropsychology*, *8*(2), 71-82. doi:10.1076/chin.8.2.71.8724

Anderson, P. J. (2001). *Measurement and development of executive function* (Doctoral dissertation, Melbourne University, Melbourne, Australia). Retrieved from https://minerva-access.unimelb.edu.au/handle/11343/35506

Anderson, P. J. (2008). Towards a developmental model of executive function. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A lifespan perspective* (pp. 3-21). New York; London: Taylor & Francis.

Anderson, P., Anderson, V., & Garth, J. (2001). Assessment and development of organizational ability: The Rey Complex Figure Organizational Strategy Score (RCF-OSS). *The Clinical Neuropsychologist*, *15*(1), 81–94. doi:10.1076/clin.15.1.81.1905

Anderson, S. W., Damasio, H., Jones, R. D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology*, *13*(6), 909–922. doi:10.1080/01688639108405107

Anderson, V. (1998). Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychological Rehabilitation, 8*(3), 319–349. doi:10.1080/713755568

Anderson, V. (2008). Attention deficits and the frontal lobes. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A lifespan perspective* (pp. 317-344). New York; London: Taylor & Francis.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, *20*(1), 385–406. doi:10.1207/S15326942DN2001_5

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

Anderson, V., Levin, H. S., & Jacobs, R. (2002). Executive functions after frontal lobe injury: A developmental perspective. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function*. (pp. 504–527). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195134971.003.0030

Anderson, V., Anderson, P., Northam, E., Taylor, H. G. (2000). Standardisation of the Contingency Naming Test for school-aged children: A measure of reactive flexibility. *Clinical Neuropsychological Assessment, 1*, 247-273.

Arbuthnott, K., & Frank, J. (2000). Trail Making Test, Part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, *22*(4), *518*–528. doi:10.1076/1380-3395(200008)22:4;1-0;FT518

Ardila, A., Rosselli, M., & Bateman, J. R. (1994). Factorial structure of cognitive activity using a neuropsychological test battery. *Behavioural Neurology*, *7*(2), 49–58. doi:10.1155/1994/131347

Arnett, P. A., Higginson, C. I., Voss, W. D., Wright, B., Bender, W. I., Wurst, J. M., & Tippin, J. M. (1999). Depressed mood in multiple sclerosis: Relationship to capacity-demanding memory and attentional functioning. *Neuropsychology, 13*(3), 434–446. doi:10.1037/0894-4105.13.3.434

Arnett, P. A., Rao, S. M., Grafman, J., Bernardin, L., Luchetta, T., Binder, J. R., & Lobeck, L. (1997). Executive functions in multiple sclerosis: An analysis of temporal ordering, semantic encoding, and planning abilities. *Neuropsychology, 11*(4), 535–544. doi:10.1037/0894-4105.11.4.535

Axelrod, B.N., Henry, R.R., & Woodward, J.L. (1992). Analysis of an abbreviated form of the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*, *6*(1), 27-31. doi:10.1080/13854049208404114

Axelrod, B.N., Jiron, C.C., & Henry, R.R. (1993). Performance of adults ages 20 to 29 on the abbreviated Wisconsin Card Sorting Test. *The Clinical Neuropsychologist, 7,* 205-209. doi:10.1080/13854049308401523

Axelrod, B.N., Paolo, A.M., & Abraham, E. (1997). Do normative data from the full WCST extend to the abbreviated WCST? *Assessment*, *4*(1), 41-46. doi:10.1177/107319119700400105

Azuma, T. (2004). Working memory and perseveration in verbal fluency. *Neuropsychology, 18*(1), 69–77. doi:10.1037/0894-4105.18.1.69

Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite reliability in structural equations modeling. *Educational and Psychological Measurement*, *55*(3), 394-406. doi:10.1177/0013164495055003003

Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 5-28. doi:10.1080/027249896392784

Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Bagozzi, R. P. (1991). Further thoughts on the validity of measures of elation, gladness, and joy. Journal of Personality and Social Psychology, 61(1), 98-104. doi:10.1037/0022-3514.61.1.98

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science, 40*(1), 8-34. doi:10.1007/s11747-011-0278-x

Banich, M. T. (2009). Executive Function: The Search for an Integrated Account. *Current Directions in Psychological Science*, *18*(2), 89–94. doi:10.1111/j.1467-8721.2009.01615.x

Banich, M. T., Burgess, G. C., Depue, B. E., Ruzic, L., Bidwell, L. C., Hitt-Laustsen, S., … Willcutt, E. G. (2009). The neural basis of sustained and transient attentional control in young adults with ADHD. *Neuropsychologia*, *47*(14), 3095–3104. doi:10.1016/j.neuropsychologia.2009.07.005

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., … Magin, R. (2000). fMRI Studies of Stroop Tasks Reveal Unique Roles of Anterior and Posterior Brain Systems in Attentional Selection. *Journal of Cognitive Neuroscience*, *12*(6), 988–1000. doi:10.1162/08989290051137521

Banich, M.T. (1997). *Neuropsychology: The neural basis of mental function*. Boston: Houghton Mifflin.

Barkley, R. A. (1988). Attention. In M. G. Tramontana & S. R. Hooper (Eds.), *Assessment issues in child neuropsychology* (pp. 145-176). New York, NY: Plenum Press.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94. doi:10.1037/0033-2909.121.1.65

Baron, I. S. (2004). *Neuropsychological evaluation of the child*. New York: Oxford University Press.

Bates, M. E., & Lemay, E. J. (2004). The d2 Test of Attention: Construct validity and extensions in scoring techniques. *Journal of the International Neuropsychological Society*, *10*(3), 392-400. doi:10.1017/S135561770410307X

Bayliss, D. M., & Roodenrys, S. (2000). Executive processing and attention deficit hyperactivity disorder: an application of the supervisory attentional system. *Developmental Neuropsychology*, *17*(2), 161–180. doi:10.1207/S15326942DN1702_02

Beauchamp, M. H., Dagher, A., Aston, J. A. D., & Doyon, J. (2003). Dynamic functional changes associated with cognitive skill learning of an adapted version of the Tower of London task. *Neuroimage*, *20*(3), 1649–1660. doi:10.1016/j.neuroimage.2003.07.003

Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian And New Zealand Journal of Public Health, 25*(5), 464-469. doi:10.1111/j.1467-842X.2001.tb00294.x.

Bennett, P. C., Ong, B., & Ponsford, J. (2005). Assessment of executive dysfunction following traumatic brain injury: Comparison of the BADS with other clinical neuropsychological measures. *Journal of the International Neuropsychological Society*, *11*(5), 606–613. doi:10.1017/S1355617705050721

Bentler, P. M., & Chih-Ping Chou. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, *16*(1), 78-117. doi:10.1177/0049124187016001004

Bentler, P. M., & Yuan, K.-H. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, *34*(2), 181–197. doi:10.1207/S15327906Mb340203

Berg, E. A. (1948). A simple objective test for measuring flexibility in thinking. *Journal of General Psychology*, *39,* 15-22. Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.proquest.com/docview/1290495773?accountid=14844

Bertola, L., Lima, M. L. C., Romano-Silva, M. A., de Moraes, E. N., Diniz, B. S., & Malloy-Diniz, L. F. (2014). Impaired generation of new subcategories and switching in a semantic verbal fluency test in older adults with mild cognitive impairment. *Frontiers in Aging Neuroscience*, *6*, 1-6. doi:10.3389/fnagi.2014.00141/full.

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development, 81*(6), 1641–1660. doi:10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180–200. doi:10.1016/j.dr.2009.05.002

Bestawros, A., Langevin, J. P., Lalonde, R., & Botez-Marquard, T. (1999). Relationship between choice reaction time and the Tower of Hanoi test. *Perceptual and Motor Skills*, *88*(2), 355–362. doi:10.2466/PMS.88.2.355-362

Bishop, D. V. M., Aamodt-Leeper, G., Creswell, C., McGurk, R., & Skuse, D. H. (2001). Individual Differences in Cognitive Planning on the Tower of Hanoi Task: Neuropsychological Maturity or Measurement Error? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(4), 551-556. doi:10.1017/S0021963001007247

Bodimeade, H. L., Whittingham, K., Lloyd, O., & Boyd, R. N. (2013). Executive function in children and adolescents with unilateral cerebral palsy. *Developmental Medicine and Child Neurology*, *55*(10), 926–933. doi:10.1111/dmcn.12195

Boone, K. B., Pontón, M. O., Gorsuch, R. L., González, J. J., & Miller, B. L. (1998). Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology*, *13*(7), 585–595. doi:10.1016/S0887-6177(97)00074-7

Bowden, S. C., & Smith, L. C. (1994). What does the Austin Maze measure? *Australian Psychologist*, *29*(1), 34-37. doi:10.1080/00050069408257317

Bowden, S. C., Dumendzic, J., Hopper, J., Kisnella, G., Clifford, C., & Tucker, A. (1992). Healthy adults' performance on the Austin Maze. *Clinical Neuropsychologist, 6(1),* 43-52. doi:10.1080/13854049208404116

Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The reliability and internal validity of the Wisconsin Card Sorting Test. *Neuropsychological Rehabilitation*, *8*(3), 243–254. doi:10.1080/713755573

Bray, R., & McDonald, S. (2010). *Austin Maze*. Sydney, NSW: Australian Society for the Study of Brain Impairment. Retrieved from https://assbi.com.au/Assessments-to-Download

Brickencamp R. 1981. Test D2: AufmerksamkeitsBelastungs-Test: Handanweisung. 7th ed. Gottingen: Verlag fur Psychologie Dr C.J. Hogrefe.

Broadbent, D. E. (1958). *Perceptions and communication.* London: Pergammon.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: a dimensional and developmental study. *Developmental Neuropsychology*, *26*(2), 571–593. doi:10.1207/s15326942dn2602_3

Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills, 84*(2), 627-661. doi:10.2466/pms.1997.84.2.627

Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: Implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology*, *5*, 1-8. doi:10.3389/fpsyg.2014.00939

Budson, A. E., & Price, B. H. (2007). Memory dysfunction in neurological practice. *Practical Neurology*, *7*(1), 42–47. Retrieved from https://pn-bmj-com.wallaby.vu.edu.au:4433/content/7/1/42

Bunce, D., & Macready, A. (2005). Processing speed, executive function, and age differences in remembering and knowing. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *58A*(1), 155–168. doi:10.1080/02724980443000197

Burgess P.W., & Shallice T. (1996b). Bizarre responses, rule detection and frontal lobe lesions. *Cortex*, *32(2),* 241-259. doi:10.1016/S0010-9452(96)80049-9

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton tests.* Thurston, Suffolk: Thames Valley Test Company.

Burgess, P.W., & Shallice T. (1996a). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34(4),* 263-272. doi:10.1016/0028-3932(95)00104-2

Campanella, F., Skrap, M., & Vallesi, A. (2016). Speed-accuracy strategy regulations in prefrontal tumor patients. *Neuropsychologia*, *82*, 1–10. doi:10.1016/j.neuropsychologia.2016.01.008

Cepeda, N. J., Blackwell, K. A., & Munakata, Y. (2013). Speed isn't everything: complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science*, *16*(2), 269–286. doi:10.1111/desc.12024

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98*(1), 67–83. doi:10.1037/0033-2909.98.1.67

Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201–216. doi:10.1016/j.acn.2007.08.010

Cheema, J. R. (2014). A review of missing data handling methods in education research. *Review of Educational Research, 84*(4), 487-508. doi:10.3102/0034654314532697

Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, *25*(3), 607–623. doi:10.1016/j.cogbrainres.2005.08.013

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting test. *Journal of Clinical and Experimental Neuropsychology*, *8*(3), 219–228. doi:10.1080/01688638608401314

Chiaravalloti, N. D., Christodoulou, C., Demaree, H. A., & DeLuca, J. (2003). Differentiating Simple Versus Complex Processing Speed: Influence on New Learning and Memory Performance. *Journal of Clinical & Experimental Neuropsychology, 25*(4), 489-501. doi:10.1076/jcen.25.4.489.13878

Cohen, R. A. (2014). *The neuropsychology of attention, 2nd ed*. New York, NY: Springer Science + Business Media. doi:10.1007/978-0-387-72639-7

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience and Biobehavioral Reviews*, *26*(2), 105–125. doi:10.1016/S0149-7634(01)00063-X

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the Unity and Diversity of the Neural Substrates of Executive Functioning. *Human Brain Mapping*, *25*(4), 409–423. doi:10.1002/hbm.20118

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–184. doi:10.1016/S0160-2896(01)00096-4

Cooper, R., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, *17*(4), 297–338. doi:10.1080/026432900380427

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain* (Doctoral dissertation, McGill University). Retrieved from https://trove.nla.gov.au/work/25065049?q&versionId=30227379

Crowe, S. F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. *Journal of Clinical Psychology*, *54*(5), 585 – 591. doi:10.1002/(SICI)1097-4679(199808)54:5<585::AID-JCLP4>3.0.CO;2-K

Crowe, S. F., Barclay, L., Brennan, S., Farkas, L., Gould, E., Katchmarsky, S., et al. (1999). The cognitive determinants of performance on the Austin Maze. *Journal of the International Neuropsychological Society, 5(1),* 1-9. doi:10.1017/S1355617799511016

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends in neurosciences, 38*(9), 571-8. doi:10.1016/j.tins.2015.07.003

Darby, D., & Darby, A. (n.d). WhiteAnt Occasional Publishing. Retrieved from http://www.brainmapping.org/WhiteAnt/

de Frias, C. M., Dixon, R. A., & Strauss, E. (2006). Structure of four executive functioning tests in healthy older adults. *Neuropsychology*, *20*(2), 206–214. doi:10.1037/0894-4105.20.2.206

de Jong, P. F., & Das-Smaal, E. A. (1993). Factor structure of standard attention tests for children: A distinction between perceptual speed and working memory. *European Journal of Psychological Assessment*, *9*(2), 94–106. Retrieved from http://wallaby.vu.edu.au:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-31898-001&site=ehost-live

De Luca, C. R., & Leventer, R. J. (2008). Developmental trajectories of executive functions across the lifespan. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A lifespan perspective* (pp. 23-56). New York; London: Taylor & Francis.

De Sonneville, L. M. J., Boringa, J. B., Reuling, I. E. W., Lazeron, R. H. C., Adèr, H. J., & Polman, C. H. (2002). Information processing characteristics in subtypes of multiple sclerosis. *Neuropsychologia*, *40*(11), 1751–1765. doi:10.1016/S0028-3932(02)00041-6 de Vries, M., & Geurts, H. M. (2012). Cognitive flexibility in ASD; Task switching with emotional faces. *Journal of Autism and Developmental Disorders*, *42*(12), 2558–2568. doi: 10.1007/s10803-012-1512-1

Debelak, R., Egle, J., Köstering, L., & Kaller, C. P. (2016). Assessment of planning ability: Psychometric analyses on the unidimensionality and construct validity of the Tower of London Task (TOL-F). *Neuropsychology*, *30*(3), 346-360. doi:10.1037/neu0000238

Della Salla, S., Gray, C., Spinnler, H., & Trivelli, C. (1998). Frontal lobe functioning in man: The riddle revisited. *Archives of Clinical Neuropsychology, 13*(8), 663–682. doi:10.1016/S0887-6177(97)00093-0

Demakis, G. J. (2003). A meta-analytic review of the sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology*, *17*(2), 255–264. doi:10.1037/0894-4105.17.2.255

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B, 39(*1), 1-38. Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsjsr&AN=edsjsr.2984875&site=eds-live

Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the Inferior Frontal Junction in Cognitive Control: Meta-Analyses of Switching and Stroop Studies. *Human Brain Mapping*, *25*(1), 22–34. doi:10.1002/hbm.20127

Diamond, A. (2000). Close Interrelation of Motor Development and Cognitive Development and of the Cerebellum and Prefrontal Cortex. *Child Development*, *71*(1), 44-56. doi:10.1111/1467-8624.00117

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function.* (pp. 466-503). New York, NY: Oxford University Press

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. doi:10.1146/annurev-psych-113011-143750

Diamond, B. J., Johnson, S. K., Kaufman, M., & Graves, L. (2008). Relationships between information processing, depression, fatigue and cognition in multiple sclerosis. *Archives of Clinical Neuropsychology*, *23*(2), 189–199. doi:10.1016/j.acn.2007.10.002

Diller, L., Ben-Yishay, Y., Gerstman, L.J., Goodkin, R., Gordon, W., & Weinberg, J. (1974). *Studies in cognition and rehabilitation in hemiplegia.* (Rehabilitation Monograph 50). New York: New York University Medical Center Institute of Rehabilitation Medicine.

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*(1), 1-17. doi:10.1186/2193-1801-2-222

Donkin, C., Little, D. R., & Houpt, J. W. (2014). Assessing the speed--accuracy trade-off effect on the capacity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1183–1202. doi:10.1037/a0035947

Dossani, R. H., Missios, S., & Nanda, A. (2015). Historical Vignette: The Legacy of Henry Molaison (1926–2008) and the Impact of His Bilateral Mesial Temporal Lobe Surgery on the Study of Human Memory. *World Neurosurgery*, *84*, 1127–1135. doi:10.1016/j.wneu.2015.04.031

Drew, M. A., Starkey, N. J., & Isler, R. B. (2009). Examining the link between information processing speed and executive functioning in multiple sclerosis. *Archives of Clinical Neuropsychology*, *24*(1), 47–58. doi:10.1093/arclin/acp007

Duke, L. M., & Kaszniak, A. W. (2000). Executive control functions in degenerative dementias: a comparative review. *Neuropsychology Review*, *10*(2), 75–99. doi:10.1023/A:1009096603879

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475–483. doi:10.1016/S0166-2236(00)01633-7

Egeth, H. E., & Yantis, S. (1997). Visual attention: control, representation, and time course. *Annual Review of Psychology*, *48*, 269–297. Retrieved from http://wallaby.vu.edu.au:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=9046562&site=ehost-live

Elston, G. N., Benavides-Piccione, R., & DeFelipe, J. (2001). The pyramidal cell in cognition: a comparative study in human and monkey. *The Journal of Neuroscience*, *21*(17), 1-5. Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-11348-001&site=eds-live

Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods*, *8*(3), 322–337. doi:10.1037/1082-989X.8.3.322

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999a). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999b). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*(3), 309–331. doi:10.1037/0096-3445.128.3.309

Eriksen, B. A.; Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non- search task. *Perception and Psychophysics, 16,* 143–149. doi:10.3758/bf03203267

Fernandez, A. L., Moroni, M. A., Carranza, J. M., Fabbro, N., & Lebowitz, B. K. (2009). Reliability of the Five-Point Test. *The Clinical Neuropsychologist*, *23*(3), 501–509. doi:10.1080/13854040802279675

Field., A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed). London: UK: Sage.

Fisk, J. E., & Sharp, C. A. (2004). Age-Related Impairment in Executive Functioning: Updating, Inhibition, Shifting, and Access. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 874–890. doi:10.1080/13803390490510680

Fodor, J. A. (1985). Prècis of The Modularity of Mind. *Behavioral and Brain Sciences*, *8*(1), 1–42. doi:10.1017/S0140525X0001921X

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1): 39-50. doi: 10.2307/3151312

Frankenhuis, W. E., & Ploeger, A. (2007). Evolutionary psychology versus Fodor: Arguments for and against the massive modularity hypothesis. *Philosophical Psychology*, *20*(6), 687–710. doi:10.1080/09515080701665904

Fry, A. F., & Hale, S. (1996). Processing Speed, Working Memory, and Fluid Intelligence: Evidence for a Developmental Cascade. *Psychological Science*, *7*(4), 237–241. doi:10.1111/j.1467-9280.1996.tb00366.x

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory and fluid intelligence in children. *Biological Psychology*, *54*(1–3), 1–34. doi:10.1016/S0301-0511(00)00051-X

Funahashi, S., & Andreau, J. M. (2013). Prefrontal cortex and neural mechanisms of executive function. *Journal of Physiology - Paris*, *107*(6), 471–482. doi:10.1016/j.jphysparis.2013.05.001

Fuster, J. M. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology*, *31*(3–5), 373–385. doi: 10.1023%2FA%3A1024190429920

Gagnon, M., Awad, N., Mertens, V. B., & Messier, C. (2003). Comparing the Rey and Taylor Complex Figures: A Test-Retest Study in Young and Older Adults. *Journal of Clinical and Experimental Neuropsychology*, *25*(6), 878–890. doi:10.1076/jcen.25.6.878.16480

Ganor-Stern, D., Seamon, J. G., & Carrasco, M. (1998). The role of attention and study time in explicit and implicit memory for unfamiliar visual stimuli. *Memory & Cognition*, *26*(6), 1187–1195. doi:10.3758/BF03201194

Gao, S., Mokhtarian, P. L., & Johnston, R. (2008). Nonnormality of Data in Structural Equation Models. *Journal of the Transportation Research Board, 2082*, 116-124. doi:10.3141/2082-14

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. doi: 10.1037/0033-2909.134.1.31

Genova, H. M., DeLuca, J., Chiaravalloti, N., & Wylie, G. (2013). The relationship between executive functioning, processing speed, and white matter integrity in multiple sclerosis. *Journal of Clinical & Experimental Neuropsychology*, *35*(6), 631–641. doi:10.1080/13803395.2013.806649

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). TEST REVIEW Behavior Rating Inventory of Executive Function. *Child Neuropsychology*, *6*(3), 235–238. doi:10.1076/chin.6.3.235.3152

Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, *8*(2), 121–137. doi:10.1076/chin.8.2.121.8727

Goel, V., & Grafman, J. (1995). Are the frontal lobes implicated in “planning” functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia*, *33*(5), 623–642. doi:10.1016/0028-3932(95)90866-P

Goel, V., Pullara, S. D., & Grafman, J. (2001). A computational model of frontal lobe dysfunction: Working memory and the Tower of Hanoi task. *Cognitive Science*, *25*(2), 287–313. doi:10.1016/S0364-0213(01)00037-4

Goldman-Rakic, P. S. (1987). Development of cortical circuitry and cognitive function. *Child Development*, *58*(3), 601–622. doi:10.2307/1130201

Goldman-Rakic, P. S. (1996). The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, *351*(1346), 1445–1453. doi:10.1098/rstb.1996.0129

Goldstein, B. E. (2008). *Cognitive Psychology Connecting Mind, research, and Everyday Experience* (2nd ed.). Belmont, CA: Thomson Wadsworth.

Goldstein, F. C., & Green, R, C. (1995). Assessment of problem solving and executive functions. In R. L. Mapou & J. Spector (Eds.*), Clinical neuropsychological Assessment: A cognitive approach* (pp.49-81). New York, NY US: Plenum Press.

Graf, P., Uttl, B., & Tuokko, H. (1995). Color- and picture-word Stroop tests: performance changes in old age. *Journal of Clinical and Experimental Neuropsychology*, *17*(3), 390–415. doi:10.1080/01688639508405132

Grant, D.A., & Berg, E.A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology, 38*(4), 404-411. doi:10.1037/h0059831

Grattan, L. M., & Eslinger, P. J. (1991). Frontal lobe damage in children and adults: A comparative review. *Developmental Neuropsychology*, *7*(3), 283–326. doi:10.1080/87565649109540496

Greenberg, L. M. (1988–2000). *T.O.V.A.® Continuous Performance Test*. Los Alamitos, CA: Universal Attention Disorders.

Greve, K. W. (2001). The WCST-64: A standardized short-form of the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*, *15*(2), 228–234. doi:10.1076/clin.15.2.228.1901

Greve, K. W., Stickle, T. R., Love, J. M., Bianchini, K. J., & Stanford, M. S. (2005). Latent structure of the Wisconsin Card Sorting Test: A confirmatory factor analytic study. *Archives of Clinical Neuropsychology*, *20*(3), 355–364. doi:10.1016/j.acn.2004.09.004

Grodzinsky, G. M., & Diamond, R. (1992). Frontal lobe functioning in boys with attention-deficit hyperactivity disorder. *Developmental Neuropsychology*, *8*(4), 427–445. doi:10.1080/87565649209540536

Gyurak, A., Goodkind, M. S., Kramer, J. H., Miller, B. L., & Levenson, R. W. (2011). Executive functions and the down-regulation and up-regulation of emotion. *Cognition & Emotion*, *26*(1), 103-118. doi:10.1080/02699931.2011.557291

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis (*7th *ed.).* Harlow, UK: Pearson Education.

Hair, J., Anderson, R. R., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis (*5th *ed.).* Upper Saddle River, NJ: Prentice-Hall.

Hale, S., & Jansen, J. (1994). Global Processing-Time Coefficients Characterize Individual and Group Differences in Cognitive Speed. *Psychological Science*, *5*(6), 384–389. doi:10.1111/j.1467-9280.1994.tb00290.x

Hancock, G. R. & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit & D. Sörbom. (Eds.). *Structural Equation Modeling: Present and Future. A Festschrift in honor of Karl Jöreskog* (pp. 195-216). Chicago: Scientific Software International.

Handley, S. J., Capon, A., Copp, C., & Harper, C. (2002). Conditional reasoning and the tower of Hanoi: The role of spatial and verbal working memory. *British Journal of Psychology*, *93*(4), 501–518. doi:10.1348/000712602761381376

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (Eds.). (1993). *Wisconsin card sorting test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology*, *20*(5), 511–528. doi:10.1037/0894-4105.20.5.511.supp

Henry, J. D., & Crawford, J. R. (2004a). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18*(2), 284–295. doi:10.1037/0894-4105.18.2.284

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*(12), 1280–1300. doi: 10.1037//0003-066X.52.12.1280

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, *45*(1–3), 73–93. doi:10.1016/S0301-0511(96)05223-4

Høie, B., Mykletun, A., Waaler, P. E., Skeidsvoll, H., & Sommerfelt, K. (2006). Executive functions and seizure-related factors in children with epilepsy in western Norway. *Developmental Medicine & Child Neurology*, *48*(6), 519–525. doi:10.1017/S0012162206001095

Howell, W. C., & Kreidler, D. L. (1963). Information processing under contradictory instructional sets. *Journal of Experimental Psychology*, *65*(1), 39–46. doi:10.1037/h0038982

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. doi:10.1080/10705519909540118

Hughes, C. (2013). Executive Function Development, Individual Differences, and Clinical Insights. In J. Rubenstein, P. Rakic & P. Rakic (Eds.), *Neural Circuit Development and Function in the Brain: Comprehensive Developmental Neuroscience* (pp. 429-445). Burlington, VT: Academic Press.

Humes, G. E., Welsh, M. C., Retzlaff, P., & Cookson, N. (1997). Towers of Hanoi and London: Reliability and Validity of Two Executive Function Tasks. *Assessment*, *4*(3), 249–257. doi:10.1177/107319119700400305

Hurks, P. P. M., Vles, J. S. H., Hendriksen, J. G. M., Kalff, A. C., Feron, F. J. M., Kroes, M., … Jolles, J. (2006). Semantic Category Fluency Versus Initial Letter Fluency Over 60 Seconds as a Measure of Automatic and Controlled Processing in Healthy School-aged Children. *Journal of Clinical and Experimental Neuropsychology*, *28*(5), 684–695. doi:10.1080/13803390590954191

Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling*, *10*(1), 128-141. doi:10.1207/S15328007SEM1001_6

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394-412. doi:10.1080/09658211003702171

Jöreskog, K. G. & Sörbom, D. (1989). *LISREL 7 User's Reference Guide.* Chicago: Scientific Software Inc.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, *36*(2), 109-133. doi.org/10.1007/BF02291393

Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model comparison approach.* New York: Harcourt Brace Jovanovich.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology Review, 17*(3), 213–233 doi:10.1007/s11065-007-9040-z

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kail, R. V., & Miller, C. A. (2006). Developmental change in processing speed: Domain specificity and stability during childhood and adolescence. *Journal of Cognition and Development*, *7*(1), 119–137. doi:10.1207/s15327647jcd0701_6

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. J*ournal of Experimental Psychology: General*, *132*(1), 47–70. doi:10.1037/0096-3445.132.1.47

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66–71. doi:10.1037/0033-2909.131.1.66

Kaplan, E., Fein, D., Morris, R., & Delis D. (1991). *WAIS-R as a neuropsychological instrument.* San Antonio, TX: The Psychological Corporation.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. doi:10.1037/bul0000160

Kay, J., & Franklin, S. (1995). Cognitive neuropsychology and assessment. In R. L. Mapou & J. Spector (Eds.), *Clinical neuropsychological Assessment: A cognitive approach* (pp.237-268). New York, NY US: Plenum Press.

Kessels, R. P. C., van den Berg, E., Ruis, C., & Brands, A. M. A. (2008). The backward span of the Corsi Block-Tapping Task and its association with the WAIS-III Digit Span. *Assessment*, *15*(4), 426–434. doi:10.1177/1073191108315611

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi Block-Tapping Task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258. doi:10.1207/S15324826AN0704_8

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York, NY: Guilford Publications. Retrieved from https://search-ebscohost-com.wallaby.vu.edu.au:4433/login.aspx?direct=true&db=cat06414a&AN=vic.b3853651&site=eds-live

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*(6), 229–235. doi:10.1016/j.tics.2007.04.005

Koechlin, E., Corrado, G., Pietrini, P., & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences of the United States of America, 97*(13), 7651-7656. doi:10.1073/pnas.130177397

Kongs, S. K., Thompson, L. L., Iverson, G. L., & Heaton, R. K. (2000). *Wisconsin Card Sorting Test-64 Card Version.* Lutz, FL: Psychological Assessment Resources.

Kortte, C. B., Horner, M. D., & Windham, W. K. (2002) The Trail Making Test, Part B: Cognitive flexibility or ability to maintain set? *Applied Neuropsychology, 9*(2), 106-109. doi:10.1207/S15324826AN0902_5

Kowalczyk, A., McDonald, S., Cranney, J., & McMahon, M. (2001). Cognitive flexibility in the normal elderly and in persons with dementia as measured by the written and oral Trail Making Tests. *Brain Impairment, 2*(1), 11-21. Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsihc&AN=edsihc.140616755108666&site=eds-live

Kuan, C. Y., Roth, K. A., Flavell, R. A., & Rakic, P. (2000). Mechanisms of programmed cell death in the developing brain. *Trends in Neurosciences*, *23*(7), 291–297. doi:10.1016/S0166-2236(00)01581-2

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*(4), 389–433. doi:10.1016/S0160-2896(05)80012-1

Lamar, M., Zonderman, A. B., & Resnick, S. (2002). Contribution of specific cognitive processes to executive functioning in an aging population. *Neuropsychology*, *16*(2), 156–162. doi:10.1037/0894-4105.16.2.156

Lamberty, G. J., Putnam, S. H., Chatel, D. M., Beliauskas, L. A., & Adams, K. S. (1994). Derived Trail Making Test indices: A preliminary report. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 7*(3), 230–234.

Landau, S. M., Garavan, H., Schumacher, E. H., & D' Esposito, M. (2007). Regional specificity and practice: Dynamic changes in object and spatial working memory. *Brain Research, 1180,* 78–89. doi:10.1016/j.brainres.2007.08.057

Langenecker, S. A., Zubieta, J. K., Young, E. A., Akil, H., & Nielson, K. A. (2007). A task to manipulate attentional load, set-shifting, and inhibitory control: Convergent validity and test–retest reliability of the Parametric Go/No-Go Test. *Journal of Clinical and Experimental Neuropsychology, 29*(8), 842-853. doi:10.1080/13803390601147611

Lee, G. P., Loring, D. W., Newell, J., & McCloskey, L. (1994). Figural fluency on the Five-Point Test: Preliminary normative and validity data. *International Neuropsychological Society Program and Abstracts*, *1,* 51.

Lee, G. P., Strauss, E., Loring, D.W., McCloskey, L., & Haworth, J. M. (1997). Sensitivity of figural fluency on the Five-Point Test to focal neurological disease. *The Clinical Neuropsychologist, 11*(1), 59–68. doi:10.1080/13854049708407030

Lehle, C., Steinhauser, M., & Hübner, R. (2009). Serial or parallel processing in dual tasks: what is more effortful? *Psychophysiology*, *46*(3), 502–509. doi:10.1111/j.1469-8986.2009.00806.x

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59. doi:10.1348/026151003321164627

Leon-Carrion, J., García-Orza, J., & Pérez-Santamaría, F. J. (2004). Development of the Inhibitory Component of the Executive Functions in Children and Adolescents. *International Journal of Neuroscience*, *114*(10), 1291–1311. doi:10.1080/00207450490476066

Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., … Fletcher, J. M. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology*, *7*(3), 377–395. doi:10.1080/87565649109540499

Levine, B., Stuss, D. T., Milberg, W. P. (1995). Concept generation: validation of a test of executive functioning in a normal aging population. *Journal of Clinical and Experimental Neuropsychology, 17(5),* 740-758. doi:10.1080/01688639508405164

Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology*, *17*(2/3), 281-297. doi:10.1080/00207598208247445

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

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.

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

Liang, L.-Y., Shewokis, P.A. and Getchell, N. (2016) Brain Activation in the Prefrontal Cortex during Motor and Cognitive Tasks in Adults. *Journal of Behavioral and Brain Science, 6*, 463- 474. doi:10.4236/jbbs.2016.612042

Little, R. J. A. & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Long, B., Spencer-Smith, M. M., Jacobs, R., Mackay, M., Leventer, R., Barnes, C., & Anderson, V. (2011). Executive function following child stroke: the impact of lesion location. *Journal of Child Neurology*, *26*(3), 279–287. doi:10.1177/0883073810380049

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Lövdén, M., Rönnlund, M., & Göran-Lars, N. (2002). Remembering and knowing in adulthood: Effects of enacted encoding and relations to processing speed. *Aging, Neuropsychology, and Cognition*, *9*(3), 184–200. doi:10.1076/anec.9.3.184.9612

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The Development of Nonverbal Working Memory and Executive Control Processes in Adolescents. *Child Development*, *76*(3), 697–712. doi:10.1111/j.1467-8624.2005.00872.x

Luria, A. (1973). *The working brain.* New York: NY: Basic Books.

Luria, A. R. (1980). *Higher cortical functions in man.* New York, NY: Basic Books.

Luria, R., & Meiran, N. (2005). Increased Control Demand Results in Serial Processing: Evidence From Dual-Task Performance. *Psychological Science*, *16*(10), 833–840. doi:10.1111/j.1467-9280.2005.01622.x

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. doi:10.1037/1082-989X.1.2.130

MacKay, D. G. (1973). Aspects of the theory of comprehension, memory and attention. *Quarterly Journal of Experimental Psychology*, *25*(1), 22-40. doi:10.1080/14640747308400320

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203. doi:10.1037/0033-2909.109.2.163

Mahurin, R. K., Velligan, D. I., Hazleton, B., Davis, J. M., Eckert, S., & Miller, A. L. (2006). Trail Making Test errors and executive function in schizophrenia and depression. *The Clinical Neuropsychologist, 20,* 271-288. https://doi.org/10.1080/13854040590947498

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32*(1), 4-18. doi:10.1111/j.1469-8986.1995.tb03400.x

Mapou, R. L. (1995). A cognitive framework for neuropsychological assessment. In R. L. Mapou & J. Spector (Eds.), *Clinical neuropsychological Assessment: A cognitive approach* (pp. 295-337). New York, NY: Plenum Press.

Mapou, R. L., & Spector, J. (1995). *Clinical neuropsychological assessment: A cognitive approach.* (3rd ed.). New York, NY: Plenum Press.

Mateer, C. A., & Williams, D. (1991). Effects of frontal lobe injury in childhood. *Developmental Neuropsychology*, *7*(3), 359–376. doi:10.1080/87565649109540498

McAuley, T., & White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology*, *108*(3), 453–468. doi:10.1016/j.jecp.2010.08.009

McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64-82. doi:10.1037/1082-989X.7.1.64

McKay, A., Lee, S., Stolwyk, R.J., & Ponsford, J.L. (2012). Comparing performance of young adults on a computer-based version of the Austin Maze and the conventional form of the test. *Brain Impairment, 13*(3), 339–346. doi:10.1017/BrImp.2012.27

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9,* 90-100. doi:10.1001/archneur.1963.00460070100010

Milner, B. (1965). Visually- guided maze learning in man: Effects of bilateral hippocampal, bilateral frontal and bilateral cerebral lesions. *Neuropsychologia, 3*(4), 317-338. doi:10.1016/0028-3932(65)90005-9

Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology Review, 2*(2), 109–145. doi:10.1007/BF01109051

Mirsky, A. F., Fantie, B. D., & Tatman, J. E (1995). Assessment of attention across the lifespan. In R. L. Mapou & J. Spector (Eds.*), Clinical neuropsychological Assessment: A cognitive approach* (pp.17-48). New York, NY: Plenum Press.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. doi: IO.II77/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex ‘‘frontal lobe’’ tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100. doi:10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*(4), 621–640. doi:10.1037/0096-3445.130.4.621

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, *21*(19), 7733–7741. Retrieved from http://wallaby.vu.edu.au:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-48202-035&site=ehost-live

Morrison, P. R., & Gates, G. R. (1988). Assessment of a microcomputer-based version of the Austin Maze. *Journal of General Psychology, 115*(3), 307-314. doi:10.1080/00221309.1988.9710567

Mosher, F., & Hornsby, J. (1966). On asking questions. In J. Bruner, et al. (Eds.), *Studies in Cognitive Growth* (pp. 86-102). New York, NY: Wiley & Sons.

Mukhopadhyay, P., Dutt, A., Kumar Das, S., Basu, A., Hazra, A., Dhibar, T., & Roy, T. (2008). Identification of neuroanatomical substrates of set-shifting ability: evidence from patients with focal brain lesions. *Progress in Brain Research*, *168*, 95–104. doi:10.1016/S0079-6123(07)68008-X

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. Davidson, R. G. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18). New York, NY: Plenum Press. doi:10 .1007/978-1-4757-0629-1_1

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*, *71*(3), 437–451. doi:10.1016/j.bandc.2009.03.005

O'Craven, K. M., Rosen, B. R., Kwong, K. K., Triesman, A., & Savoy, R. L. (1997). Voluntary attention modulates fMRI activity in human MT/MST. *Neuron, 18*(4), 591–598. doi:10.1016/S0896-6273(00)80300-1

Olson, E. A., & Luciana, M. (2008). The development of prefrontal cortex functions in adolescence: Theoretical models and a possible dissociation of dorsal versus ventral subregions. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed.). (pp. 575–590). Cambridge, MA: MIT Press.

Oosterman, J. M., Wijers, M., & Kessels, R. P. C. (2013). Planning or something else? Examining neuropsychological predictors of Zoo Map performance. *Applied Neuropsychology: Adult*, *20*(2), 103–109. doi:10.1080/09084282.2012.670150

Orr, J. M., Sackett, P. R., & Dubois, C. L. (1991). Outlier detection and treatment in I/O psychology: A survey of researcher beliefs and an empirical illustration. *Personnel Psychology, 44*(3), 473-486. doi:10.1111/j.1744-6570.1991.tb02401.x

Osterrieth, P. A. (1944). Le test de copie d'une figure complex: Contribution a l'étude de la perception et de la mémoire. *Archives de Psychologie, 30,* 286–356.

Owens, E. M., Denney, D. R., & Lynch, S. G. (2013). Difficulties in planning among patients with multiple sclerosis: A relative consequence of deficits in information processing speed. *Journal of the International Neuropsychological Society, 19*(5), 613–620. doi:10.1017/S1355617713000155

Pachella, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pachella, R. G., Fisher, D. F., & Karsh, R. (1968). Absolute judgments in speeded tasks: Quantification of the trade-off between speed and accuracy. *Psychonomic Science, 12*(6), 225–226. doi:10.3758/BF03331282

Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *Journal of Clinical and Experimental Neuropsychology, 33*(4), 456–470. doi:10.1080/13803395.2010.533157

Pallant, J. F. (2007). *SPSS survival manual: a step by step guide to data analysis using SPSS for Windows (version 15)* (3rd ed.). Crows Nest, NSW: Allen & Unwin.

Paolo, A.M., Axelrod, B.N., & Troster, A. (1996). Test-retest stability of the Wisconsin Card Sorting Test. *Assessment, 3*(2), 137-143. doi:10.1177/107319119600300205

Partington, J. E., & Leiter, R. G. (1949). Partington's Pathway Test. *The Psychological Service Center Bulletin*, *1*, 9–20.

Pashler, H. (1994a). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin*, *116*(2), 220–244. doi:10.1037/0033-2909.116.2.220

Pashler, H. (1994b). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 330–342. doi:10.1037/0096-1523.20.2.330

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Child Psychology & Psychiatry & Allied Disciplines*, *37*(1), 51–87. doi:10.1111/j.1469-7610.1996.tb01380.x

Peter, B., Matsushita, M., & Raskind, W. H. (2011). Global processing speed in children with low reading ability and in children and adults with typical reading ability: Exploratory factor analytic models. *Journal of Speech, Language, and Hearing Research*, *54*(3), 885–899. doi:10.1044/1092-4388(2010/10-0135)

Pietrzak, R. H., Maruff, P., Mayes, L. C., Roman, S. A., Sosa, J. A., & Snyder, P. J. (2008). An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. *Archives of Clinical Neuropsychology*, *23*(4), 433–445. doi:10.1016/j.acn.2008.03.002

Pineda, D. A., & Merchan, V. (2003). Executive function in young Colombian adults. *International Journal of Neuroscience*, *113*(3), 397–410. doi:10.1080/00207450390162164

Pires, L., Moura, O., Guerrini, C., Buekenhout, I., Simões, M. R., & Leitão, J. (2018). Confirmatory Factor Analysis of Neurocognitive Measures in Healthy Young Adults: The Relation of Executive Functions with Other Neurocognitive Functions. *Archives of Clinical Neuropsychology, 34*(3), 350-365. doi:10.1093/arclin/acy040

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The Neural Correlates of Motor Skill Automaticity. *The Journal of Neuroscience*, *25*(22), 5356–5364. doi:10.1523/JNEUROSCI.3880-04.2005

Porteus, S. D. (1965). *Porteus Maze Test: Fifty years' application*. Palo Alto: California Pacific Books.

Posner, M. I. & Petersen, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25– 42. doi:10.1146/annurev.ne.13.030190.000325

Purdy, M. H. (2016). Executive functions: Theory, assessment, and treatment. In M. L. Kimbarow (Ed.), *Cognitive communication disorders* (2nd ed.). (pp. 83–128). San Diego, CA: Plural Publishing. Retrieved from http://wallaby.vu.edu.au:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2014-52160-003&site=ehost-live

Ranti, C., Chatham, C. H., & Badre, D. (2015). Parallel temporal dynamics in hierarchical cognitive control. *Cognition*, *142*, 205–229. doi:10.1016/j.cognition.2015.05.003

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*(2), 173-184. doi:10.1177/01466216970212006

Raykov, T., & Widaman, K. F. (1995). Issues in Applied Structural Equation Modeling Research. *Structural Equation Modeling: A Multidisciplinary Journal*, *2*(4), 289–318. doi:10.1080/10705519509540017

Regard, M. (1981). *Cognitive rigidity and flexibility: A neuropsychological study* (Unpublished doctoral dissertation). University of Victoria.

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production of verbal and nonverbal fluency tasks. *Perceptual and Motor Skills*, *55*(3), 839–844. doi:10.2466/pms.1982.55.3.839

Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology*, *19*(5), 393–394. doi:10.1037/h0044509

Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie*, *28*, 286–340.

Ríos, M., Periáñez , J. A. , & Muñoz-Céspedes, J. M. (2004). Attentional control and slowness of information processing after severe traumatic brain injury. *Brain Injury*, *18*(3), 257-272. doi:10.1080/02699050310001617442

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. A. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society*, *4*(5), 474–490. doi:10.1017/S1355617798455073

Robertson, I.H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention*. Bury St. Edmunds: Thames Valley Test Company.

Romine, C. B., & Reynolds, C. R. (2005). A Model of the Development of Frontal Lobe Functioning: Findings from a Meta-Analysis. *Applied Neuropsychology*, *12*(4), 190–201. doi:10.1207/s15324826an1204_2

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211–227. doi:10.1037/0096-3445.126.3.211

Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., et al. (2002). Executive control function: A review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neurosciences*, *14*, 377–405. doi:10.1176/jnp.14.4.377

Ruff, R. M. (1988). *Ruff Figural Fluency Test Professional Manual.* Odessa, FL: Psychological Assessment Resources Inc.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. doi:10.1037/0033-295X.103.3.403

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top–down meets bottom–up. *Brain Research Reviews*, *35*(2), 146–160. doi:10.1016/S0165-0173(01)00044-3

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research, 8*(1), 3-15. doi:10.1191/096228099671525676

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*(2), 147-177. doi:10.1037/1082-989X.7.2.147

Schmidt, C. S. M., Schumacher, L. V., Römer, P., Leonhart, R., Beume, L., Martin, M., … Kaller, C. P. (2017). Are semantic and phonological fluency based on the same or distinct sets of cognitive processes? Insights from factor analyses in healthy adults and stroke patients. *Neuropsychologia*, *99,* 148–155. doi:10.1016/j.neuropsychologia.2017.02.019

Schnirman, G. M., Welsh, M. C., & Retzlaff, P. D. (1998). Development of the Tower of London-Revised. *Assessment*, *5(4),* 355-360. doi:10.1177/107319119800500404

Schretlen, D. (1997). *Brief Test of Attention professional manual*. Odessa, FL: Psychological Assessment Resources.

Schubert, A.-L., Hagemann, D., & Frischkorn, G. T. (2017). Is general intelligence little more than the speed of higher-order processing? *Journal of Experimental Psychology: General*, *146*(10), 1498–1512. doi:10.1037/xge0000325

Shallice, T. (1979). Case study approach in neuropsychological research. *Journal of Clinical Neuropsychology, 1*(3), 183-211. doi:10.1080/01688637908414450

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of The Royal Society of London. Series B, Biological Sciences, 298*(1089), 199–209. doi:10.1098/rstb.1982.0082

Shallice, T., & Burgess, P. W. (1996). The domain of supervisory processes and temporal organization of behavior. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 351,* 1405-1412. doi:10.1098/rstb.1996.0124

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

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., … DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology*, *34*(5), 585–602. doi:10.1007/s10802-006-9037-8

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*, 1-10. doi:10.3389/fpsyg.2014.00772

Shiffrin, R. M., & Schneider, W. (1984). Automatic and controlled processing revisited. *Psychological Review*, *91*(2), 269–276. doi:10.1037/0033-295X.91.2.269

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81*, 174-176. doi:10.1037/h0027448

Slocum-Gori, S. L., Zumbo, B. D., Michalos, A. C., & Diener, E. (2009). A note on the dimensionality of quality of life scales: An illustration with the Satisfaction with Life Scale (SWLS). *Social Indicators Research*, *92*(3), 489-496. doi:10.1007/s11205-008-9303-y

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition*, *22*(1), 1–13. doi:10.3758/BF03202756

Solso, R. L., MacLin, O. H., MacLin, K. M. (2014). *Cognitive Psychology*. (8th ed.). Harlow, UK: Pearson Education.

Span, M. M., Ridderinkhof, K. R., & van der Molen, M. W. (2004). Age-related changes in the efficiency of cognitive processing across the life span. *Acta Psychologica*, *117*(2), 155–183. doi:10.1016/j.actpsy.2004.05.005

Stolwyk, R. J., Lee, S., McKay, A., & Ponsford, J. L. (2013). Exploring what the Austin maze measures: A comparison across conventional and computer versions. *Brain Impairment*, *14*(2), 243-252. doi:10.1017/BrImp.2013.23

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

Strauss, M. E., Thompson, P., Adams, N. L., Redline, S., & Burant, C. (2000). Evaluation of a model of attention with confirmatory factor analysis. *Neuropsychology, 14*(2), 201–208. doi:10.1037/0894-4105.14.2.201

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology, 18*(6), 643–662. doi:10.1037/h0054651

Stuss, D. T. (2006). Frontal lobes and attention: Processes and networks, fractionation and integrations. *Journal of the International Neuropsychological Society, 12,* 261-271. doi:10.1017/S1355617706060358

Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society, 17*(5), 759–765. doi:10.1017/S1355617711000695

Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research, 63,* (3-4), 289-298. doi:10.1007/s004269900007

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin, 95*(1), 3–28. doi:10.1037/0033-2909.95.1.3

Stuss, D. T., Alexander, M. P., Floden, D., Binns, M. A., Levine, B., McIntosh, A. R., … Hevenor, S. J. (2002). Fractionation and localization of distinct frontal lobe processes: Evidence from focal lesions in humans. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function.* (pp. 392–407). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195134971.003.0025

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., … Izukawa, D. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*(4), 388–402. doi:10.1016/S0028-3932(99)00093-7

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of the New York Academy of Sciences, 769,* 191-212.

Stuss, D. T., Toth, J. P., Franchi, D., Alexander, M. P., Tipper, S., & Craik, F. I. (1999). Dissociation of attentional processes in patients with focal frontal and posterior lesions. *Neuropsychologia*, *37*(9), 1005–1027. doi:10.1016/S0028-3932(98)00158-4

SüB, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence*, *30*(3), 261. doi:10.1016/S0160-2896(01)00100-3

Swensson, R. G. (1972). The elusive tradeoff: Speed vs accuracy in visual discrimination tasks. *Perception & Psychophysics*, *12*(1–A), 16–32. doi:10.3758/BF03212837

Swiercinsky, D. P., & Hallenbeck, C. E. (1975). A factorial approach to neuropsychological assessment. *Journal of Clinical Psychology*, *31*(4), 610–618. doi:10.1002/1097-4679(197510)31:4<610::AID-JCLP2270310407>3.0.CO;2-J

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Pearson/Allyn & Bacon.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson Education.

Taylor, L. B. (1991). Scoring criteria for the ROCF. In O. Spreen., & E. Strauss, *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.

Testa, R., Bennett, P., & Ponsford, J. (2012). Factor analysis of nineteen executive function tests in a healthy adult population. *Archives of Clinical Neuropsychology*, *27*(2), 213–224. doi:10.1093/arclin/acr112

Thomas, E., Reeve, R., Fredrickson, A., & Maruff, P. (2011). Spatial memory and executive functions in children. *Child Neuropsychology*, *17*(6), 599–615. doi:10.1080/09297049.2011.567980

Thurstone, L. L. (1938). *Primary Mental Abilities.* Chicago, IL: University of Chicago Press.

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 3–18. doi:10.1037/0096-1523.29.1.3

Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin, 20,* 12-16. doi:10.1093/oxfordjournals.bmb.a070274

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and healthy adults. *Neuropsychology*, *11,* 138–146. doi:10.1037/0894-4105.11.1.138

Tucker, A., Kinsella, G., Gawith, M., & Harrison, G. (1987). Performance on the Austin Maze: Steps towards normative data. *Australian Psychologist*, *22(3)*, 353-359. doi:10.1080/00050068708259536

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *The Quarterly Journal of Experimental Psychology*, *64*(3), 447–466. doi:10.1080/17470218.2010.505292

Uttl, B., & Graf, P. (1997). Color-Word Stroop test performance across the adult life span. *Journal of Clinical And Experimental Neuropsychology*, *19*(3), 405–420. doi:10.1080/01688639708403869

Verbruggen, F., & Logan, G. D. (2008). Automatic and controlled response inhibition: Associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General*, *137*(4), 649-672. doi:10.1037/a0013170

Vestberg, T., Reinebo, G., Maurex, L., Ingvar, M., & Petrovic, P. (2017). Core executive functions are associated with success in young elite soccer players. *PLoS ONE*, *12*(2), 1–13. doi:10.1371/journal.pone.0170845

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, *3*(4), 255–274. doi:10.3758/CABN.3.4.255

Walsh, K. (1994). *Neuropsychology: A clinical approach* (3rd ed.). Edinburgh, U.K.: Churchill Livingstone.

Walsh, K. W. (1978). *Neuropsychology: A clinical approach.* New York, NY: Churchill Livingston.

Walsh, K. W., & Darby, A. (1994). *Neurosychology: A Clinical Approach* (3rd ed.). Edinburgh: Churchill Livingston.

Walsh, K.W. (1991). *Understanding brain damage: A primer of neuropsychological evaluation* (2nd ed.). Edinburgh, U.K: Churchill Livingstone.

Wechsler, D. (1997a). *WAIS-III Administration and Scoring manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997b). *Technical Manual for the Wechsler Adult Intelligence Test* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation.

Welsh, M. C., & Pennington, B. F. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental Neuropsychology*, *4*(3), 199–230. doi:10.1080/87565648809540405

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, *7*(2), 131–149. doi:10.1080/87565649109540483

Welsh, M. C., Satterlee-Cartmell, T., & Stine, M. (1999). Towers of Hanoi and London: Contributions of working memory and inhibition to performance. *Brain and Cognition*, *41*(2), 231–242. doi:10.1006/brcg.1999.1123

Welsh, M., Cicerello, A., Cuneo, K., & Brennan, M. (1995). Error and temporal patterns in Tower of Hanoi performance: Cognitive mechanisms and individual differences. *Journal of General Psychology*, *122*(1), 69–81. doi:10.1080/00221309.1995.9921223

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult*, *23*(1), 29–34. doi:10.1080/23279095.2015.1004574

Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefficient for multidimensional measures: A pedagogical illustration. *Review of Psychology*, *21*(2), 111-121. Retrieved from https://hrcak.srce.hr/147109

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575–587. doi:10.1037/0012-1649.44.2.575

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

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, *73*(6), 913-934. doi:10.1177/0013164413495237

Yerys, B. E., Wallace, G. L., Harrison, B., Celano, M. J., Giedd, J. N., & Kenworthy, L. E. (2009). Set-shifting in children with autism spectrum disorders: Reversal shifting deficits on the

Intradimensional/Extradimensional Shift Test correlate with repetitive behaviors. *Autism*, *13*(5), 523–538. doi:10.1177/1362361309335716

Yun, J.-Y., Lee, D. Y., Seo, E. H., Choo, I. H., Park, S. Y., Kim, S. G., & Woo, J. I. (2011). Neural Correlates of Stroop Performance in Alzheimer's Disease: A FDG-PET Study. *Dementia and Geriatric Cognitive Disorders Extra*, *1*(1), 190–201. doi:10.1159/000329517

Zainudin, A. (2012). *A handbook on SEM: Structural equation modelling using amos graphics* (4th ed.). Kelantan: University Technology MARA Press.

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 445–469). Malden, MA: Blackwell.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, *1*(2), 198–226. doi:10.1037/1089-2680.1.2.198

Zhang, W. (2015). A structural equation modeling approach to factors that influence farmers' behaviour and behavioural intentions towards water policy changes (Doctoral dissertation, University of Lethbridge, Alberta, Canada). Retrieved from http://wallaby.vu.edu.au:2048/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-99180-125&site=ehost-live

Ziegler, M., & Hagemann, D. (2015). Testing the unidimensionality of items: Pitfalls and loopholes. *European Journal of Psychological Assessment*, *31*(4), 231-237. doi:10.1027/1015-5759/a000309

Zillmer, E. A., & Spiers, M. V, & Culbertson, W. C. (2008). *Principles of neuropsychology*. (2nd ed.). Belmont, CA, US: Wadsworth/Thomson Learning.

Zoccatelli, G., Beltramello, A., Alessandrini, F., Pizzini, F. B., & Tassinari, G. (2010). Word and position interference in stroop tasks: a behavioral and fMRI study. *Experimental Brain Research, 207*(1–2), 139–147. doi:10.1007/s00221-010-2433-x

Zygmont, C., & Smith, M. R. (2014). Robust factor analysis in the presence of normality violations, missing data, and outliers: Empirical questions and possible solutions. *The Quantitative Methods for Psychology, 10*(1), 40-55. doi: 10.20982/tqmp.10.1.p040

Appendix A

Please note, the following table is applicable as a reference guide for all tables in the proceeding appendices, and should be referred to in the absence of a Table legend.

Table 14

Summary of Total Variables used in Initial Data Reduction, Scoring Methods and Purported Measurement

Test	Variable name	Scoring method (T=Traditional, N= Non-traditional, D=Derived) formula, and purported test interpretation (H= higher score, better performance L=lower score, better performance)	Purported measure	Reference
Attention tasks				
Trails (TMT)	1. TMT-A 2. TMT-B 3. TMT B-A	1. Total time (T) (L) 2. Total time (T) (L) 3. B time- A time difference score (T) (L)	1a) Visual search, speed, attention 1b) Selective attention, self-regulation 2a) Visual search, speed, attention, divided attention 2b) divided attention and mental flexibility 3. Purer measure of more	1+2a) Lamberty et al. (1994); Strauss et al. (2006) 1+2b) Anderson (2001) 3. Lamberty et al. (1994); Strauss et al. (2006)

			complex divided attention, mental flexibility, and speed	
d2	4. d2 H 5. d2 M 6. d2 FA 7. d2 total error 8. d2CONC (concentrate performance)	4. Total correctly cancelled (T) (H) 5. Omission errors (T) (L) 6. Commission errors (T) (L) 7. Omission + commission errors (T) (L) 8. Hits-FA (N) (H)	4-7. Sustained, selective and scanning attention, processing speed 8. Speed and accuracy components of concentration performance	4-7. Brickencamp (1981) 8. Bates & Lemay (2004)
Stroop	9. Stroop	9. Colours time/Dots time (T) (L)	9. Inhibition, suppressing habitual response, cognitive control and flexibility, selective attention, behavioural inhibition	9. Strauss et al. (2006); Stroop (1935); Graf et al. (1995) for the interference score
Map Search 2	10. Map Search	10. Correctly circled symbols after 2 minutes (T) (H)	10. Selective attention	10-14. Robertson et al. (1994)
Visual Elevator timing score	11. VE	11. Only for correct responses: time/switches (T) (L)	11. Attentional switching and cognitive flexibility	

Elevator Counting with Reversal	12. ECR	12. Total correct out of 10 (T) (H)	12. Auditory verbal WM
Telephone Search while Counting	13. TSC-dual task decrement 14. TSC-E	13. Time per target weighted for accuracy of tone counting – time per parget of Telephone Search from previous sub test (T) (L) 14. Time/correctly circles symbols= time per target score. Then time per target score/ proportion of correctly counted tones= Time per target weighted for accuracy of tone counting (N) (L)	13. Divided attention, and ability complete complex tasks and multitask, however was initially meant as a measure of divided attention. Previous subtest of TS is used as 'motor control task' where the dual task decrement score accounts and controls for individual variation in processing or psychomotor speed 14. Divided attention, and ability complete complex tasks, multitask, however was initially meant as a measure of divided attention.
Span tasks			

Block-span	15. BS-Fwd 16. BS-Bk	15. Total trials correct (T) (H) 16. Total trials correct (T) (H)	15. Spatial attention and concentration 16. Spatial WM	15-16. Brunetti et al. (2014); Smyth and Scholey, (1994); Wechsler, (1997b)
Digit-span	17. DS-Fwd 18. DS-Bk	17. Total trials correct (T)((H) 18. Total trials correct (T) (H)	17. Efficiency of attention and concentration 18. Working memory capacity/mental tracking	17-18. Wechsler, (1997b); Lezak et al., (2004)
Fluency tasks				
Verbal Fluency	19. FAS-15 (phonemic) 20. FAS-AV 21. FAS total 22. FAS cluster size 23. FAS switches 24. Animals-15 (semantic) 25. Animals-AV 26. Animals total	19.Total admissible words at 15 seconds (N) (H) 20. Words at 30,45,60/3 (D) (H) 21. Total admissible words (T) (H) 22. Successively generated words that begin with same first two letters (counted by beginning with second word in each cluster (descriptive) (N) (H) 23. Number of transitions between clusters (descriptive) (N) (H) 24. Total admissible words at 15 seconds (N) (H)	19. Task initiation/maintenance, automatic processing 20. Controlled processing, WM, attentional control 21. Fluency, word production, lexical access speed, Executive Function, vocabulary retrieval (e.g., automatic activation from a cue, self-monitoring, inhibition, generation of cues) 22. Memory and word storage 23. Cognitive flexibility, shifting, Executive Function	19. Delis et al. (2001) as reported in Strauss et al. (2006); Hurks et al. (2006) 20. Present study, adapted from last 15 second interval from Hurks et al. (2006). 21. Fisk and Sharp, (2004); Shao et al. (2014); Rosen & Engle (1997)

		25. Words at 30,45,60/3 (D) (H) 26. Total admissible words (T) (H)	24-26. As above	22. Troyer et al. (1997) 23. Bertola et al. (2014); Troyer et al. (1997) 24. Delis et al. (2001), Hurks et al. (2006) 25. Present study, adapted from last 15 second interval from Hurks et al. (2006) 26. Fisk and Sharp, (2004); Shao et al. (2014); Rosen & Engle (1997)
5-point	27. 5-point total 28. 5-point-1st 29. 5-point-AV 30. 5-point % perseveration errors (PPE)	27. Total Unique Designs (UD) (T) (H) 28. Total UD at 1st minute (N) (H) 29. Average UD of last 2 trials (D) (H)	27. Nonverbal fluency, executive control 28. Task initiation/maintenance, automatic processing 29. Controlled nonverbal processing, WM	27. Regard et al. (1982) 28-29. Present study 30. Lee et al. (1997)

		30. [(Perseverative errors/total unique designs) x100] (T) (L)	30. Perseverative behaviour	
Planning tasks				
Austin Maze (AM)	31. AM total errors 32. AM total time	31. Cumulative errors to criterion; 2 error free trials) (T) (L) 32. Total time to criterion (2 error free trials (T) (L)	31a) Spatial ability, WM, visuospatial learning b) EF c) planning 32. Speed element of planning and visuospatial learning	31a) Crowe et al. (1999) 31b) Walsh (1978) 31c) Bray & McDonald (2010); Milner (1965); Walsh & Darby, (1994) 32. Present study
Tower of Hanoi (TOH)	33. TOH Bishop 34. TOH total time 35. TOH total moves 36. TOH total correct 37. TOH residual	33. Highest level of task successfully completed, with a ½ point added if both tasks at the level were passed (in terms of minimum number of moves) i.e passed both 3 moves, 1 4 move, and both 5 moves= 5.5 (1/2 point added if both completed at this level) (N) (H)	33a) Planning 34-36. Planning, inhibition, processing speed, problem-solving, working memory, visuospatial memory and problem solving 37. Self-regulation	33a) Bishop et al. (2001) 34-36. Arnett et al., (1997); Shallice, (1982); Goel & Grafman, (1995); Miyake et al. (2000); Goel, et al. (2001)

		34. Total time to complete all 13 trials (T) (L) 35. Total moves to complete all 13 trials (T) (L) 36. Total correct out of 13 who made it in minimum moves (T) (H) 37. How many moves over the minimum: Total moves-minimum moves (81) (D) (L)		37. Present study
Rey Osterrieth Complex Figure Task (ROCFT)	38. ROCFT copy time 39. ROCFT copy raw score 40. ROCFT derived score 41. ROCFT ORG	38. Total time to complete (N) (L) 39. Copy accuracy score 0-2 points for each section as determined by placement (T) (H) 40. Copy time/ copy raw accuracy (D) (L) 41. Level of organisation (7- excellent, 1- unrecognizable) (T) (H)	38. Amount of information retained over time 39. Accuracy, planning, visual-constructional ability, organisation, problem solving, motor functioning 40. Efficiency of planning and organisation 41. Organisation	38. Strauss et al. (2006) 39. Meyers & Meyers, (1995b); as reported in Strauss et al. (2006); 40. Present study 41. Anderson (2001)
Rule Shift	42. Rule Shift	42. Total time to correctly identified 20 cards based on rule (T) (L)	42. Shifting, keeping track of rules	42. Wilson et al. (1996)
Key Search	43. Key Search time	43. Total time to complete (N) (L) 44. Raw accuracy score (T) (H)	43. Organisation time 44. Efficiency of search, plan,	43-44. Wilson et al. (1996)

	44. Key Search raw score 45. Key Search derived score	45. Time/ copy raw accuracy (D) (L)	monitor performance, 45. Efficiency of organized search including their speed of processing to plan efficiently	45. Present study
Zoo Map	46. Zoo Map planning time 47. Zoo Map total time 48. Zoo Map raw score 49. Zoo Map raw score inverse 50. Zoo Map derived score	46. Planning time before pen to paper (T) (L) 47. Total time to complete (total time-planning time for independence of observations) (D) (L) 48. Correct places visited- total errors (T) (H) 49. Semi derived: Constant of 9 added to raw to score to remove negative value (those who made more errors than correct sequence) (D) (H) 50. Total time / by raw score (which includes sequence score and errors) constant score of 9 was added to raw score to remove the negative	46. Initial planning ability 47. Total planning ability 48. Planning ability, modify performance based on feedback 49. Planning ability, modify performance based on feedback 50. Efficiency of planning ability, including speed and errors used to assess self-monitoring and regulation	46-48. Wilson et al. (1996) 49-50. Present study, although based on Wilson et al.'s (1996) scoring

		value (those who made more errors than correct sequence) (D) (L)		
Picture Arrangement	51. PA	51. Raw score as per manual instructions (T) (H)	51. Concept formation, reasoning, processing speed, non-verbal reasoning and organisation, sequential	51. Lezak et al. (2004); Wechsler (1997a)
Wisconsin Card Sorting Test (WCST)	52. WCST total correct 53. WCST total errors 54. WCST P errors 55. WCST non P errors 56. WCST P responses 57. WCST categories completed	52. Responses that match the sorting principle (T) (H) 53. Sum of P errors and non P errors (incorrect sorting principle) (T) (L) 54. Perseverative errors, where a person persists on sorting to an incorrect rule (T) (L) 55. Incorrect sorting rule (T) (L) 56. Responses that match the perseverated to principle are scores as P responses regardless if they are correct or incorrect (T) (L) 57. Categories completed i.e 10 consecutive correct matches per criterion (min 0 to max 6) (T) (H)	52. Overall the WCST is a measure of EF	52-57. Heaton et al. (1993)

Appendix B

Table 15

Test Administration Order

Version A	Version B	Version C	Version D
n=34	*n*=33	*n*=33	*n*=33
WASI	WASI	WASI	WASI
TMT-A	TOH	5 point	Blocks
TMT-B	TMT-A	Stroop	Digits
FAS	TMT-B	TMT-A	TEA
TOH	PA	TMT-B	AM
d2	FAS	ROCFT	PA
Blocks	WCST	FAS	FAS
Digits	Blocks	d2	**BREAK**
Stroop	Digits	Animals	TOH
ROCFT	BADS	BADS	Animals
5 point	**BREAK**	**BREAK**	WCST
Animals	TEA	Blocks	TMT-A
AM	Stroop	Digits	TMT-B

BREAK	AM	TEA	Stroop
PA	ROCFT	AM	ROCFT
WCST	5 point	PA	5 point
BADS	d2	TOH	d2
TEA	Animals	WCST	BADS

Note: BADS order always= Rule Shift, Key Search, Zoo Map, TEA order always= Map Search, *EC, VE, ECR, TS, TSC*

Appendix C

Table 16

Excluded Tests

No.	Test	Proposed domain	Failed criterion
1	Random Number Generation (Brugger, 1997)	Working Memory	Norms available for a limited age range Difficult to locate in public domain
2	Porteus Maze (Porteus, 1965)	Planning	Austin Maze was preferred over others that had very similar task demands
3	Six Elements (Wilson et al., 1996)	Planning, organisation, monitoring behaviour	Did not have outcome measures of speed and accuracy
4	Hayling sentence completion task	Response inhibition, fluid ability (Strauss et al., 2006)	Not available in public domain

	(Burgess & Shallice, 1996a)		
5	Brixton spatial anticipation test (Burgess & Shallice, 1996b)	Rule derivation, shifting, fluid ability (Strauss et al., 2006)	Not available in public domain
6	Twenty Questions (Mosher & Hornsby, 1996)	Feedback utilization, problem solving, categorical processing (Strauss et al., 2006)	Not available in public domain- was available under D-KEFS Organisation did not have access
7	Contingency naming task (Anderson, Anderson, Northam, & Taylor, 2000)	Inhibition, mental flexibility (Strauss et al., 2006)	No normative data was available for adults and scoring guidelines were for children Limited upward psychometric properties for middle to later adulthood

8	Concept generation (Levine, Stuss & Milberg, 1995)	Inhibition (Levine, Stuss, & Milberg, 1995)	WCST was preferred over others that had very similar task demands Time restraints
9	Similarities (Wechsler, 1997a)	Abstract verbal reasoning	Only needed 2 subtests, Block Design and Vocabulary are most commonly used. Only used as a measure of estimated IQ not as part of analysis
10	Tower of London-R (Schnirman, Welsh, & Retzlaff, 1998)	Planning, WM, inhibition	Difficult to distinguish between healthy and controls in adults (Baron, 2004) TOL more so a measure of simple planning, when rather the TOH is more of a complex measure of planning and targeted towards response inhibition.

			Pegs were simple, one sized, only distinguished by different colours, whereas TOH had different sized discs thereby making this more difficult, in order to tap into EF
11	Design Fluency (Ruff, 1998)	Response (design) generation, Executive Function (Strauss et al., 2006)	Couldn't be used as alternate version of fluency No specific material available Instruction are vague 15 minute administration time Inter rater reliability poor
12	Cognitive estimates (Shallice & Evans, 1978)	Problem solving (Strauss et al., 2006)	No normative data was available Did not have outcome measures of speed and accuracy

			Did not produce effective outcome measures for the present study. Simple tasks that did not really tap into the complexities noted as EF
13	Test of Variables of Attention (TOVA; Greenberg, 1988-2000)	Sustained inhibition, response speed, attention	Not available in the public domain (computerized test) Time constraints (20 minute administration time)
14	Sentence arrangement (Kaplan, et al., 1991)	Sequential reasoning and syntactically correct constructions (Kaplan et al., 1991)	Did not produce effective outcome measures for the present study. Simple tasks that did not really tap into the complexities noted as EF
15	Modified Six Slements (from the BADS test, Wilson et al., 1996)	Organisation, planning, monitoring behaviour (Strauss et al., 2006)	Did not have outcome measures of speed and accuracy Did not produce effective outcome measures for the present study (similar to Cognitive estimates). Simple tasks that did not really tap into the complexities noted as EF

16	Paced Auditory Serial Addition Test (part of the Brief Test of Attention (Schretlen, 1997)	Sustained Attention WM, information processing, speed, divided attention (Strauss et al., 2006)	Time constraints (20 minute administration time) Present study only needed one major battery of attention tests (consisting of same administration time) went with the TEA instead as wanted selective and controlled attention.
17	Action program (Wilson et al., 1996)	Organisation, planning (Strauss et al., 2006), problem solving	Did not have outcome measures of speed and accuracy Did not produce effective outcome measures for the present study (similar to Cognitive estimates and Modified Six Elements). Simple tasks that did not really tap into the complexities noted as EF
18	Temporal judgment (Wilson et al., 1996)	Organisation, planning (Strauss et al., 2006).	Did not have outcomes measures of speed and accuracy.

			Did not produce effective outcome measures for the present study (similar to Cognitive estimates, Modified Six Elements, and Action Program) Simple tasks that did not really tap into the complexities noted as EF
19	Ruff figural fluency (Ruff, 1998)	Planning (Strauss et al., 2006)	Not available in public domain Less sensitive to unique designs Multiple trials redundant Age differences in young and old Education and IQ impacted scores
20	Simon task (1969)	Inhibition and selection of sensory information	Available on computer program PEBL limited availability of norms

			Similar to the Stroop test
21	Flanker test (Eriksen & Eriksen, 1974)	Interference and response inhibition	Available on computer program PEBL limited availability of norms
22	Go/no go task (Verbruggen, & Logan, 2008)	Response inhibition	Available on computer program PEBL
23	N back task, and dual n back task (Jaeggi, Buschkuehl, Perrig, & Meier, 2010)	WM	More of a WM measure Low reliability, therefore used for research (Jaeggi et al., 2010)

24	Line /letter cancellation (Diller, Ben-Yishay, Gerstman, Goodkin, Gordon, & Weinberg, 1974)	Visual inattention, spatial neglect	Whilst this does have a shared attentional component, the neglect aspect is not really pertinent for the present study Qualitative interpretation
25	Symbol cancellation (Diller et al., 1974)	Visual inattention, neglect	Whilst this does have a shared attentional component, the neglect aspect is not really pertinent for the present study Qualitative interpretation
26	Elevator Counting with Distraction	Sustained attention	Are not mediated by the anterior network, and are better indicators of posterior attention
27	Lottery	Sustained attention	Are not mediated by the anterior network, and are better indicators of posterior attention

Appendix D

Flyers

Do you like puzzles?

Are you creative?

Are you good at completing brain teasers?

Do you want to test the range of skills you have?

If you answered YES to any of these questions, we are looking for you! If you are aged between 18-80 years old, we would like you to participate in our study relating to how well you can process a range of cognitive skills ranging in task difficulty.

We are looking for adults who would like to complete a range of fun easy and complex tasks, to help us with our research into Executive Function and how we can better understand this 'general thinking' construct.

What does participation involve?

You will complete some tasks that assess general 'thinking' skills reflecting the following:

- Visual tasks, such as block tapping
- Verbal tasks, such as naming words
- Drawing tasks
- Maze puzzle tasks
- Organisation tasks

If you would like more information about the study or would like to participate, please contact **Dr Michelle Ball** on **9919 2536** or email **michelle.balll@vu.edu.au**, **Dr Emra Suleyman** on **9919 2397** or email **emra.suleyman@vu.edu.au, or Jessica Burlak on 0411 575 176** or email **jessica.burlak@vu.edu.au**

Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au	Adult Executive Function Project Dr Michelle Ball: 9919 2536; michelle.ball@vu.edu.au

Appendix E

Information to participants involved in research

You have been invited to participate in a research project

You are invited to participate in a research project entitled *"Exploring the factor structure of Executive Functioning: Establishing a psychometrically robust framework for the assessment of Executive Function in adults"*

This project is being conducted by Dr. Michelle Ball and Dr. Emra Suleyman from the College of Arts, at Victoria University, together with PhD candidates Jessica Burlak, and Adam Bromage, and Psychology Honours students, Scott Mc Donald, Sarah Hill and Jamiee Roach.

Project explanation

Executive functions (EF) are those abilities we use to allow us to undertake complex tasks. This entails the "hard stuff" like planning, organising and monitoring our own behaviour, just to name a few. Psychologists have been interested in EF for many decades, but because they drive complex behaviour it has been difficult for us to develop tests that accurately measure them. To explain – when we think about planning, don't we also need organisation skills to plan effectively? Then, in carrying out our plan, don't we need to monitor our behaviour in order to keep ourselves on track in relation to the plan? These questions become really important when someone has a problem with EF. Unless we have tests that can effectively separate out the different EF (or acknowledge overlap) it is hard for clinicians to tell whether the person is having trouble with just one, or all of these functions.

Our project aims to explore how well current tests are able to assess and differentiate the various EF skills. Currently there is controversy about the role of attention in EF, and we are also hoping to inform the debate about this by adding tests of attention. We will then use the knowledge gained about all of the tests to inform a model of EF that includes attention which has only been used in children previously.

This knowledge will help us identify the process of EF, so that efficient assessments of EF for adults can take place, in turn helping identify if someone is having difficulty, and in what way. This will also help Psychologists with making appropriate rehabilitation recommendations to those that have difficulty.

What will I be asked to do?

You will be asked to complete a range of computer and pen to paper based tests of thinking and EF at whichever location is most convenient for you between a quiet room in your home, or at a campus of Victoria University. The tests will include a variety of things including manual skills such as arranging blocks and discs, or drawing tasks, connecting the dots and lastly, verbal tasks. You will also be asked to complete a few tasks that provide an estimate of your intelligence. Altogether testing will take about 3 hours, and as you can imagine, it will be quite tiring. There will be lots and lots of tests administered, but each one should only take a few minutes to complete. To help to minimise your fatigue we would like to give you a choice about how the testing is carried out. We can either arrange two 90 minute testing sessions on the one day, with a 1 hour break in the middle, or we can hold the two 90 minute sessions on separate days (but with no more than 7 days between each session). Whichever suits you best. Additionally we have a few questionnaires about your thinking skills and your personality we would like you to complete in your own time and deliver to us at the testing session.

Please note that those who have been diagnosed with a developmental disorder (e.g., ADHD, Autism Spectrum Disorder, Dyslexia), neurological disorder (e.g., Cerebral Palsy, Muscular Dystrophy etc.), Psychological Disorder (e.g., Depression) and whose estimated IQ is previously known to be significantly below that of their same-age peers will not be eligible to participate in the study.

How will the information I give be used?

Your anonymous data will be used for preparation of written journal articles, research theses and/or conference presentations. The information gained will be used for different theses as follows:

- Jessica Burlak – PhD exploring the factor structure of EF (and using MOST of the data)
- Adam Bromage – PhD developing a new framework contributing to the theory underlying EF
- Scott McDonald – Honours investigating how well EF test results apply to real world situations
- Sarah Hill – Honours investigating whether the way people respond to social situations will predict their EF ability
- Jamiee Roach – Honours investigating personality and EF

Only collated data will be reported, and no identifying information about any individual will be used in the preparation of any publications. If any member of the research team is known to you, that person will not have access to your individual data and another member of the team will complete all test administration and scoring procedures to protect your privacy.

What are the potential risks of participating in this project?

The process of undertaking cognitive skills assessment can be intimidating and participants may experience some anxiety over this. A certain amount of test anxiety is a normal feeling, and researchers will try to ensure that this process is as fun and easy as possible. If you feel too anxious then you can stop the assessment at any time with no negative consequences to yourself.

Should you choose to take no further part in the study all documentation relating to your personal details and assessments will be shredded. Registered psychologist Dr. Jenny Sharples, at Victoria University has agreed to be contacted should you need to discuss any psychological issues arising from this study. She has agreed to discuss treatment options and arrange referral to appropriate services if necessary. She can be contacted on 9919 4448, or jenny.sharples@vu.edu.au.

We also understand that we are asking you to take part in a lot of testing, and acknowledge that this may make you feel tired. That is the reason we have asked to meet you for two sessions, instead of just one. Furthermore, we will offer you a break half way through each testing session should you require it. Should you continue to feel tired then you can withdraw your participation at any time with no penalty to yourself and all confidential records and personal details about you will be shredded.

What will I gain from participating?

Although we can promise no direct benefit to you, you will be participating in research that hopes to contribute significantly to the understanding of EF. It is hoped that we can use this information to improve diagnosis and rehabilitation for people who suffer from difficulty with these important skills.

Who is conducting the study?

This study is being conducted by the College of Arts at Victoria University by Dr Michelle Ball, Dr Emra Suleyman, PhD candidates Jessica Burlak and Adam Bromage and Honours in Psychology students Scott McDonald, Sarah Hill and Jamiee Roach. Contact details of senior members of the research team are provided blow. Please feel free to contact any member of the team should you have any queries about your participation. Note that Jessica and Adam are the primary contacts if you would like to register interest in taking part.

Michelle Ball	Emra Suleyman
michelle.ball@vu.edu.au	emra.suleyman@vu.edu.au
9919 2536	9919 2397
Jessica Burlak	Adam Bromage
Jessica.burlak@vu.edu.au	adam.bromage@vu.edu.au
0411 575 176	0414 229 314

If you have any queries or complaints about the way you have been treated, you may contact the Research Ethics and Biosafety Manager, Victoria University Human Research Ethics Committee, Victoria University, PO Box 14428, Melbourne, VIC, 8001 or phone (03) 9919 4148.

Appendix F

Consent form for participants involved in research

INFORMATION TO PARTICIPANTS:

You have been invited to participate in a research project entitled *“Exploring the factor structure of Executive Functioning: Establishing a psychometrically robust framework for the assessment of Executive Function in adults"* that is being conducted by Dr Michelle Ball and Dr Emra Suleyman from the College of Arts, at Victoria University, together with PhD candidates Jessica Burlak, and Adam Bromage, and Psychology Honours students, Scott Mc Donald, Sarah Hill and Jamiee Roach.

The purpose of this study has been explained to me in the Information to Participants’ form and any questions I have about the study have been answered by a member of the research team.

CERTIFICATION BY SUBJECT

I, (participants name)____________________________

of (suburb)_____________________________________ certify that I am at least 18 years old and that I am voluntarily giving my consent to participate in the study named above. I also confirm that I have no pre-existing or current neurological, psychological or developmental disorder.

I certify that the objectives of the study, together with any risks and safeguards associated with the procedures listed hereunder to be carried out in the research, have been fully explained to me by a member of the research team, and that I understand that I can withdraw from this study at any time and that this withdrawal will not jeopardise me in any way.

I have been informed that the confidentiality of the information I provide will be safeguarded. Participation in this project will be anonymous, so my identity will remain confidential. I have been informed that my information will be stored confidentially by the researchers at Victoria University (VU).

I freely consent to participation involving the below mentioned procedures:

- Completing a series of questionnaires that will be provided to me and I will return to the researchers at a testing session

- Taking part in an extensive series of assessments of my EF, either in a quiet room of my own home or at a VU campus. I understand that this testing will take up to 3 hours to complete and that I will be offered several breaks.

Signed:

Date:

Any queries about your participation in this project may be directed to the researchers

Dr Michelle Ball	Dr Emra Suleyman	Jessica Burlak
03 9919 2536	03 9919 2397	0411 575 176
Michelle.ball@vu.edu.au	emra.suleyman@vu.edu.au	Jessica.burlak@vu.edu.au

If you have any queries or complaints about the way you have been treated, you may contact the Ethics Secretary, Victoria University Human Research Ethics Committee, Office for Research, Victoria University, PO Box 14428, Melbourne, VIC, 8001, email Researchethics@vu.edu.au or phone (03) 9919 4781 or 4461.

Appendix G

Table 17

Summary of Assumptions, Various Purported Treatment Methods, those Utilised, and the Outcomes of Assumptions for all Steps of the Research Design

Assumption	Rules of thumb/treatment methods purported in the literature	Treatment method used/outcome in current study
Univariate outliers	Remove if legitimate (Orr et al., 1991).	All were used when necessary.
	Remove if extreme (Orr et al., 1991).	
	Remove to ensure there are honest parameters in a data set (Judd, & McClelland, 1989).	

	Removing based on extremity does not suffice, therefore truncation is best (Tabachnick & Fidell, 2007). Replace with the mean (Tabachnick & Fidell, 2007).	
Univariate normality	Acceptable range of cut off points of 3.00 for skewness and 10 for kurtosis (Kline, 2015).	All values fell within acceptable ranges. Of those slightly kurtotic, still did not exceed cut off.
Missing data acceptable cut off	5% (Schafer, 1999). 10% (Bennett, 2001). 15-20% (Enders, 2003).	*Variable selection and reduction phase*: <5% for most, 7.5% for WCST[a] variables, 22.6 for TOH[b], 15.8% for AM[c].

	20% (Wolf et al., 2013).	*Confirmation of latent constructs:* 8.8% for TOH, 9.0% for AM. *Predictive validity and interrelationships between constructs:* all constructs were equal at *N*=113.
Missing data treatment methods	*Conventional approaches*; Listwise deletion, Pairwise deletion, Dummy Variable Adjustment, Imputation. *Model based approaches;* Maximum Likelihood Estimation (MLE): Expectation Maximization (EM), Direct Maximum Likelihood (ML) (Dempster et al., 1997) all must fall under the multivariate normality and the MAR assumption. *Other methods;* Multiple Imputation (MI), Monte Carlo algorithm.	Multivariate normality and MAR[d] were met, therefore the model based approach EM[e] was employed. WCST was not estimated in the variable selection and reduction phase as estimated score did not make theoretical sense (P and non P errors must add up to total errors, which then must be the difference score to total correct which did not happen when estimated). This was re-evaluated and estimated after data reduction included WCST total correct for congeneric modelling, and

corrections were made if necessary (cannot obtain a higher score from the maximum).

TOH and AM were not estimated during the variable selection and reduction phase as they were above the 5% cut off, however were re-evaluated for the congeneric model analysis. To avoid specification error, although reducing sample size is a limitation, data were deleted (20 cases) and MVA[f] was calculated on TOH and. This resulted in the sample dropping to *N*=113 to confirm the Goal Setting construct, and the remaining constructs left at *N*=133.

Given the unequal sample size between constructs it was essential to rectify in order to confirm the

		interrelationships amongst them. Therefore the entire sample was reduced removing the same 20 cases from TOH and AM from prior to meet the Goal Setting construct leaving a total of *N*=113 for the entire study from here on in.
Multivariate normality	SEM[g] is sensitive to deviations from Multivariate Normality (MVN), however is often impractical to examine (Kline, 2015). Meeting univariate normality on every variable often suffices (Kline, 2015).	Given MLE[h] was used and is robust to these violations and meeting univariate normality suffices which was met in the current study, Mardia's coefficient was calculated as a precaution. Of the small cases that demonstration potential threats, biases were small confidence lies in the present data to be free from threats of multivariate non-normality, indicating MLE as the appropriate estimation procedure to run.

	Maximum Likelihood Estimation (MLE) robust to violations of normality (Hu & Bentler, 1999).	
	Mardia's coefficient and the critical ratio (c.r) < 1.96.	
Multivariate outliers	Mahalanobis distance <1 (Allen & Bennett, 2012; Tabachnick & Fidell, 2013).	Of the few that were slightly violated, Cook's D were <1.
Factors determining sample size	Model misspecification.	All theoretically viable variables were included in this comprehensive approach (as much as possible), therefore not violated.

	Model size: 10, 5, 3, and 2 respondents per parameter (Bagozzi and Yi, 2012; Bentler and Chou, 1987; Hair et al., 1998).	Model size was met at 10 participants per parameter for all besides Goal Setting construct which was met at 5 per parameter.
	Departures from normality.	
	Estimation procedure.	There were no extreme departures from multivariate normality, therefore the estimation procedure MLE could be applied that is robust to deviations from appropriate sample size.
Unidimensionality	All directions consistent, and meeting reliability and validity cut offs.	Psychological constructs are rarely unidimensional (Slocum-Gori, et al., 2009) as they are made up of multiple facets that contribute to performance, therefore

		difficult to satisfy (Brunner & Süβ, 2005 as cited in Widhiarso & Ravand, 2014). The same premise applies to EF, therefore interpret with caution.
Construct validity	Good fit according to overall fit indices.	All constructs demonstrated good overall fit indices. Overall model fit usually suffices.
	Indicators range between .5-.8 (Bagozzi, 1991; Fornell & Larcker, 1981; Hair, et al., 2014; Nunnally & Bernstein, 1994 as cited in Hancock Mueller, 2001).	
Convergent validity	Indicator variables <.*p*=.05.	All indicators were significant, and ones that were not were removed.

	Indicator loadings between .5-.7 (Hair et al., 2014), and individual reliability of an indicator .4-.7 (Hair et al., 1998; Zainudin, 2012).	Not all indicators met these cut off scores. However EF[i] is made up of multiple facets therefore not an unexpected flaw, therefore interpret with caution.
AVE[j]	>.5.	As above, as the AVE is calculated from the indicator variables.
Discriminant validity	Free of redundancy (i.e., no modification indices).	Warranted if theoretically supported. In the instances where this did occur, they were theoretically valid.

[a=]Wisconsin Card Sorting Test [b=]Tower of Hanoi [c=]Austin Maze [d=]Missing At Random [e=]Expectation Maximisation [f=]Missing Value Analysis [g=]Structural Equation Modelling [h=]Maximum Likelihood Estimation [i=]Executive Function [j=]Average Variance Explained

Appendix H

Percent of missing data

7.5% missing for WCST variables, 82.7% for WCST attempt (however was not used in any analyses, was used purely for accurate estimation methods) followed by 6.0% for d2 total error, 5.3% d2CONC, M and FA, 4.5% for d2 H, 2.3% for Key Search time, 1.5% for ECR, and .8% for TMT B-A, ROCFT derived score and total time, FAS cluster size and switches, 5-point and its variants apart from 5-point-AV which was 1.5%, .8% for TSC timing score, TSC-dual task, TSC-E, Key Search derived score, and Zoo Map derived and raw score respectively. TOH had 22.6% and the AM had 15.7%.

Appendix I

Table 18

Data Reduction Approach Presenting Included and Excluded Variables and Descriptive Statistics of all 57 Variables Including the M, SD, Minimum and Maximum scores and Possible Range

	Variable	*N*	Number of correlations *r*=.3	*M*	*SD*	Min	Max	Possible range
1	**VE (T)**	133	17	3.77	0.86	2.00	7.30	-
2	**BS-Bk (T)**	133	16	6.55	1.84	3.00	12.00	0+
3	d2 H (T)	133	15	187.11	43.80	40.00	294.00	300
4	**d2CONC (N)**	133	15	185.12	44.49	28.00	294.00	-58-300
5	**AM total time (T)**	112	12	304.22	78.92	143.34	483.07	0-∞
6	**TMT-B total time (T)**	133	12	57.18	17.80	20.06	114.50	0-∞
7	**ECR (T)**	133	11	7.90	2.23	2.00	10.00	0-10

8	TMT B-A (T) difference score	133	10	33.96	15.57	6.91	84.82	-
9	**TOH total time (T)**	103	10	306.43	111.67	128.73	727.80	0-∞
10	**DS-Bk (T)**	133	9	6.08	2.66	1.00	14.00	0+
11	**DS-Fwd (T)**	133	9	6.74	2.31	2.00	14.00	0+
12	**PA (raw score) (T)**	133	9	13.86	3.73	4.00	20.00	0-22
13	FAS-15 (N)	133	8	16.74	4.26	7.00	30.00	0+
14	**FAS total (T)**	133	7	42.42	10.92	16.00	71.00	0+
15	WCST categories complete (T)	123	7	3.67	1.50	0.00	6.00	0-6
16	**Rule Shift (T)**	133	7	27.94	6.78	15.26	56.50	0-∞
17	FAS-AV (D)	133	6	8.56	2.74	2.67	14.67	0+
18	**Animals total (T)**	133	6	26.62	5.31	13.00	41.00	0+

19	AM total error (T)	112	5		65.44	31.04	21.00	159.00	0+
20	**BS-Fwd (T)**	133	5		6.77	1.91	2.00	11.00	0+
21	**5-point total (T)**	133	5		34.57	8.03	9.00	53.00	0+
22	**TMT-A total time (T)**	133	5		23.22	6.91	10.56	49.81	0-∞
23	FAS switches (N)	133	5		26.54	7.77	8.00	45.00	0+
24	Animals-AV (D)	133	5		5.33	1.47	1.67	9.00	0+
25	**WCST total correct (T)**	123	5		48.70	9.13	15.00	59.00	0-64
26	WCST total errors (T)	123	5		15.30	9.13	5.00	49.00	0-64
27	FAS cluster size (N)	133	4		9.95	5.52	1.00	27.00	0+
28	WCST non P errors (T)	123	4		7.20	5.42	1.00	26.00	0-64

29	WCST P responses (T)	123	4	8.85	5.78	3.00	32.00	-
30	5-point-1st (N)	133	3	16.89	4.79	5.00	27.00	0+
31	WCST P errors (T)	123	3	8.02	4.73	3.00	26.00	-
32	Zoo Map total time (accurate) (D)	133	3	81.63	51.60	13.65	265.87	0-∞
33	ROCFT copy time (N)	133	2	131.49	40.08	51.12	268.03	0-330s
34	**ROCFT derived score (D)**	133	2	4.47	1.49	2.17	11.17	-
35	**Stroop (T)**	133	2	1.68	0.43	0.99	2.83	-
36	TOH total correct (T)	103	2	6.43	2.28	2.00	12.00	0-13

37	**Map Search (T)**	133	2	68.72	8.65	43.00	80.00	0-80
38	**TSC-E (N)**	133	2	2.95	0.88	1.70	5.30	-
39	d2 FA (T)	133	1	2.03	2.91	0.00	12.00	358
40	5-point-AV (D)	133	1	8.84	2.72	2.00	19.00	0+
41	TOH total moves (T)	103	1	133.14	30.56	83.00	224.00	0+
42	TOH residual (D)	103	1	52.14	30.56	2.00	143.00	0+
43	Animals-15 (N)	133	1	10.65	2.29	5.00	16.00	0+
44	**Zoo Map derived score (D)**	133	1	7.26	6.37	0.85	23.60	-
45	d2 M (T)	133	0	19.85	15.92	1.00	83.00	-
46	d2 total error (T)	133	0	21.84	16.43	1.00	80.00	-
47	5-point % perseveration (T)	133	0	5.90	6.32	0.00	21.88	0+

48	ROCFT copy score (T)	133	0	29.84	3.95	17.00	36.00	0-36
49	ROCFT org (T)	133	0	5.05	1.24	2.00	7.00	0-7
50	TOH Bishop (N)	103	0	5.77	1.85	0.00	9.50	13.5
51	TSC-dual task (without outliers) (T)	133	0	0.35	0.70	-1.20	2.30	-
52	Key Search time (N)	133	0	62.10	45.71	8.63	219.14	0-∞
53	Key Search raw score (T)	133	0	12.89	3.58	3.00	16.00	0-16
54	**Key Search derived score (D)**	133	0	5.38	4.34	0.54	16.43	-
55	Zoo Map planning time (N)	133	0	59.16	55.98	1.00	213.00	0-∞

56	Zoo Map raw score (T)	133	0	4.46	4.12	-8.00	8.00	-8-8
57	Zoo Map inversion (D)	133	0	13.46	4.12	1.00	17.00	1-17

Note: All variables in bold were retained for further analysis. All span tasks are all trials correct. (T)=Traditional scoring approach. (N)=Non-traditional scoring. (D)=Derived: scoring procedures created by researcher.

Appendix J

Table 19

Summary of Correlation Analysis

Variable	*N*	Statistical: Inter correlations *r*=>.3	Theoretically supported if failed statistically? Y/N	Inter correlations *r*=>.3 or [inter correlations *r*=>.25] if not supported statistically
VE	133	17		TMT-A, TMT-B, TMT B-A, DS-Bk, d2 H, d2CONC, Animals total, Animals-AV, FAS-15, FAS total, FAS-AV, FAS clusters, 5-point total, ECR, Rule Shift, TOH time, AM total time
BS-Bk	133	16		TMT-A, TMT-B, TMT B-A, DS-Bk, d2 H, d2CONC, Animals total, ECR, TOH time, AM total time, AM error, WCST total correct, WCST

			total errors, WCST P responses, WCST P errors, WCST cat comp
d2CONC	133	15	TMT-A, TMT-B, TMT B-A, ROCFT copy, ROCFT derived, DS-Bk, BS-Bk, DS-Fwd, 5-point-1st, 5-point total, VE, ECR, Rule Shift, AM total error, AM total time
d2 H	133	15	TMT-A, TMT-B, TMT B-A, ROCFT copy, ROCFT derived, DS-Bk, BS-Bk, DS-Fwd, 5-point-1st, 5-point total, VE, ECR, Rule Shift, AM total error, AM total time
TMT B total time	133	12	BS-Fwd, DS-Bk, BS-Bk, DS-Fwd, d2 H, d2CONC, Map Search, VE, ECR, TSC-E, Rule Shift, Zoo Map total time (accurate)

AM total time	112	12	BS-Fwd, BS-Bk, d2 H, d2CONC, 5-point-1st, 5-point total, VE, ECR, WCST total correct, WCST total errors, WCST non P errors, WCST cat comp
ECR	133	11	TMT-B, TMT B-A, PA, BS-Bk, d2 H, d2CONC, VE, Rule Shift, TOH time, AM total error, AM total time
TOH total time	103	10	BS-Bk, VE, ECR, Rule Shift, WCST cat comp, WCST total error, WCST total correct, WCST P errors, WCST non P errors, WCST p responses
TMT B-A difference score	133	10	BS-F, DS-Bk, BS-Bk, DS-Fwd, d2 H, d2CONC, VE, ECR, Zoo Map total time (accurate), Zoo Map derived

PA (raw score)	133	9	ECR, TOH total correct, AM total error, WCST total correct, WCST total error, WCST P responses, WCST P errors, WCST non P errors, WCST cat comp
DS-Fwd	133	9	TMT-B, TMT B-A, Stroop, d2 H, d2CONC, FAS-15, FAS total, FAS-AV, FAS switches
DS-Bk	133	9	TMT-B, TMT B-A, BS-Fwd, BS-Bk, d2 H, d2CONC, FAS total, FAS-AV, VE
FAS-15	133	8	DS-Fwd, Animals-15, Animals total, FAS clusters, FAS switches, 5-point-AV, 5-point total, VE
Rule Shift	133	7	TMT-B, d2 H, d2CONC, VE, ECR, Zoo Map total time, TOH total time

WCST categories complete	123	7	PA, BS-Bk, TOH moves, TOH total time, TOH residual, TOH total correct, AM total time
FAS total	133	7	DS-Bk, DS-Fwd, Animals total, Animals-AV, FAS clusters, FAS switches, VE
Animals total	133	6	BS-Bk, FAS-15, FAS-AV, FAS total, FAS switches, VE
FAS-AV	133	6	DS-Bk, DS-Fwd, Animals total, FAS clusters, FAS switches, VE
WCST total errors	123	5	PA, BS-Bk, Animals-AV, TOH total time, AM total time
WCST total correct	123	5	PA, BS-Bk, Animals-AV, TOH total time, AM total time

Animals-AV	133	5	FAS total, VE, WCST total correct, WCST total error, WCST non P error
FAS switches	133	5	DS-Fwd, Animals total, FAS-15, FAS-AV, FAS total
TMT-A total time	133	5	d2 H, d2CONC, VE, WCST P responses, Map Search
5-point total	133	5	d2 H, d2CONC, VE, FAS-15, AM total time
BS-Fwd	133	5	TMT-B, TMT B-A, DS-Bk, d2 FA, AM total time
AM total error	112	5	PA, BS-Bk, d2 H, d2CONC, ECR
WCST P responses	123	4	TMT-A, PA, BS-Bk, TOH total time

WCST non P errors	123	4		PA, Animals-AV, TOH total time, AM total time
FAS clusters	133	4		FAS-15, FAS-AV, FAS total, VE
Zoo Map total time (accurate)	133	3		TMT-B, TMT B-A, Rule Shift
WCST P errors	123	3		PA, BS-Bk, TOH total time
5-point-1st	133	3		d2 H, d2CONC, AM total time
TSC-E	133	2	Y	TMT-B, Stroop
Map Search	133	2	Y	TMT-A, TMT-B
TOH total correct	103	2		WCST cat comp, PA
Stroop	133	2	Y	DS-Fwd, TSC-E

ROCFT derived score	133	2	Y	d2 H, d2CONC [WCST cat comp, AM total error, TOH total correct TMT-A]
ROCFT copy time	133	2		d2 H, d2CONC
Zoo Map derived score	133	1	Y	TMT B-A [TMT-B]
Animals-15	133	1		FAS-15
TOH residual	103	1		WCST cat comp
TOH total moves	103	1		WCST cat complete
5-point-AV	133	1		FAS-15

d2 FA	133	1		BS-Fwd
Zoo Map inversion	133	0		
Zoo Map raw score	133	0		
Zoo Map planning time	133	0		
Key Search derived score	133	0	Y	[AM total time]

Key Search raw score	133	0
Key Search time	133	0
TSC-dual task (without outliers)	133	0
TOH Bishop	103	0
ROCFT org	133	0
ROCFT copy score	133	0
5-point % perseveration	133	0
d2 total error	133	0

d2 M	133	0