Reconceptualising the Nature of Executive Functioning: Introducing the Hierarchical Demand Model (HDM)

Adam Carl Bromage

This thesis is submitted in partial fulfillment of the requirements of the degree Doctor of Philosophy (Psychology)

School of Psychology

College of Health and Biomedicine

Victoria University

Melbourne

Australia

JULY 2020

ABSTRACT

Over the past 20 years, many investigations have sought to define the construct of Executive Functioning (EF) via the fragmentation of Executive Function skillsets. This approach has served useful for the demarcation and detailed account of complex cognitive functions that can manifest during various clinical pathologies. Yet, theoretical accounts remain heterogeneous in these representations, and neuropsychological measures of their purported constituents are critiqued for measurement impurity and poor ecological validity. This study aimed to reconceptualise traditional scoring approaches of a variety of EF tests through the lens of cognitive control theory, via a 3-study design. Performance on 13 different EF tests were assessed in 105 participants (M_{age}= 30.00, *SD* = 7.11). Study 1 sought to develop, apply, and preliminary test a Demand Classification System (DCS). The DCS comprised of 16 classifications of demand across a criterion of complexity (Abstraction, Contextual Stability, Action Rules, Instruction and Rules, Dual-Task Requirements) and novelty (Automaticity, Schematic, and Episodic). Confirmatory Factor Analysis (CFA) demonstrated that *within* these tests, performance variance could be attributed to the nature of its internal demands for complexity and novelty. Using congeneric modelling, Study 2 revealed that complexity-novelty classifications accounted for a significant amount of unique shared variance *between* tests. Factor weighted performance scores were created for each level of demand identified by Studies 1 and 2. Finally, Study 3 aimed to assess the relationships between performance at varying levels of demand using Structural Equation Modelling. The results supported the existence of hierarchically contingent relationships that align with existing neurological mechanisms proposed for cognitive control. The collective outcome of these analyses is the offering of The Hierarchical Demand Model (HDM) which proposes that in order to successfully engage in a goal-directed task the individual must *recognise* what is known, *appraise* what is required, and *reconcile* the difference to formulate an effective response. The HDM and its agent the DCS, collectively serves to operationalise known influencers of demand to the neuropsychological testing environment. This has significant implications not only for the future use of neuropsychological tools as singular measures of EF, but also how these measures can be better utilised for the assessment of overall higher-order cognitive ability. This approach urges the recognition of synergy between cognitive control and EF, and the duality of their influence over the execution of controlled behaviour in response to demand. This project offers not only a methodical system whereby this synergistic approach can be successfully applied, but also a framework that is able to account for the nature of human engagement at various levels of complexity and novelty across a hierarchical continuum of demand.

DECLARATION

I, Adam Carl Bromage, declare that the PhD thesis entitled "Reconceptualising the Nature of Executive Functioning: Introducing the Hierarchical Demand Model (HDM)" 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.

Signed: ..[redacted]

Date: 21st July 2020

DEDICATION

For my husband Eddie.

Thank you for bringing sunshine to each and every day.

ACKNOWLEDGMENTS

It is only at the very end that you can truly begin to appreciate the journey. Like many others, this thesis represents moments in time where friendships have formed, lessons have been learned, and achievements made. I feel lucky to present this thesis with bitter-sweet feelings, as I farewell a creation that I love. A feeling that is only possible due to the support that I have received along the way.

I would firstly like to thank the individuals who participated in this study for committing their precious time and energy. Your selflessness and willingness to contribute to this project reflects the very best qualities of human nature.

Thank you to my parents, family and friends who have tirelessly shown pride, interest, patience and fierce support over the years that this thesis has developed. Thank you to Bree Jones for your continuous support, wisdom, and diligent eye at all stages of my academic journey.

To my husband Eddie, thank you for believing in me, inspiring me, consoling me, and loving me. Thank you for your unwavering patience and support. It has been many years now since we danced around the room with excitement at my first acceptance into university. Now, as this final chapter closes, it is time for us to dance together again, at last.

To Assoc. Prof Michelle Ball, thank you for your supervision, friendship, guidance and for helping to unpack my often initially outlandish ideas. Thank you for encouraging deep thought and consideration when it was needed most, so that what emerged held appropriate strength and value.

Finally, To Dr Emra Suleyman, you have continuously challenged my abilities at every step of this journey. You have nurtured the growth of not only a researcher, but also shaped the values and perspectives of an individual. My love for how this project has eventuated is owed considerably to your selfless belief and dedication towards ensuring that my journey is one filled with personal pride. Thank you for this everlasting gift. Many years ago, in your office, when voicing my frustrations with Executive Functioning, you challenged me to do something about it, and here it is.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

ABBREVIATIONS

ADF	Asymptotic Distribution-Free
AIC	Akaike Information Criterion
AM	Austin Maze
BD	Block Design Test
C&F	Complex and Familiar
C&N	Complex and Novel
CE	Central Executive
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
DCS	Demand Classification System
DF	Degrees of Freedom
DLPFC	Dorsolateral Prefrontal Cortex
DMN	Default Mode Network
ECR	Elevator Counting with Reversal
EF	Executive Functioning
EFA	Exploratory Factor Analysis
EFs	Executive Functions
EMN	Extrinsic Mode Network
FSIQ-2	Full Scale IQ- 2
GDC	Global Demand Classification
GFI	Goodness-of-Fit Index
HDM	Hierarchical Demand Model
ICN	Intrinsic Control Network
LPFC	Lateral Pre-Frontal Cortex
M-distance	Mahalanobis distance
MD	Multiple Demand System
MFST	Multi Feature Sorting Task
MLE	Maximum Likelihood Estimation
PFC	Pre-Frontal Cortex
QSM	Quasi-simplex model
RLPFC	Rostrolateral PFC
RMSEA	Root Mean Square Error Of Approximation

SD	Standard Deviation
S&F	Simple and Familiar
S&N	Simple and Novel
SAS	Supervisory Attention System
SEM	Structural Equation Modelling
SRMR	Standardised Root Mean Square Residual
S→R	Stimulus-Response
TEA	Test of Everyday Attention
TMT	Trail Making Test
TOH	Tower of Hanoi
TOL	Tower of London
VIF	Variance Inflation Factor
WAIS	Wechsler Adult Intelligence Scale
WASI	Wechsler Abbreviated Scale of Intelligence – Second Edition
WCST	Wisconsin Card Sorting Test
WM	Working Memory
WMC	Working Memory Capacity

PREAMBLE

Human beings are capable of complex thought and behaviour more than any other species. Modern man is charged with the responsibility of handling numerous complex tasks, often simultaneously, and sometimes without warning. To accomplish these tasks, behaviour is required to be controlled, refined, and well-suited to the current set of demands, with the expectation that the task is completed both accurately and efficiently. Our cognitive systems have evolved to support these expectations. During early development, pattern recognition and schema formation are some of the earliest complex cognitive processes to emerge. Pattern recognition serves to enable efficiency of task responding, so that we can be adaptive, flexible, and engage with a multitude of tasks to exert control and mastery. As our pattern recognition and schemas become more sophisticated, this allows us to allocate all of our cognitive resources to the novel aspects of task demands to further engage in new goal-directed and controlled behaviour. Our ability to succeed during a familiar environment hinges on our capacity to convert initial novel experiences into familiar and learned knowledge that can then later be recalled and applied to future behaviour. Whereas our capacity to successfully engage in behaviour that is both controlled and adaptive to novel task demands requires a sophisticated cognitive system that is still not fully understood.

Amongst all of the known cognitive systems, Executive Functioning (EF) features prominently for enabling the engagement of goal-directed behaviour (Anderson et al., 2011; Barkley, 2012; Best & Miller, 2010; Dawson & Guare, 2010; Lezak et al., 2012; Luria, 1976; Miller & Cohen, 2001; Oosterlaan et al., 2005; Welsh & Pennington, 1988) to respond adaptively to novel (Anderson et al., 2011; Barkley, 2012; Norman & Shallice, 1986; Snyder et al., 2015) and complex environment demands (Diamond, 2013; García-Madruga et al., 2016; Hughes & Graham, 2002; Jurado & Rosselli, 2007; Lezak et al., 2012). Decades of theoretical, clinical and neuroanatomical research has amassed to demonstrate the contribution and importance of EF as a higher-order human ability (Baddeley, 2007; Friedman & Miyake, 2016; Lezak et al., 2012; Stuss, 2011). However, unlike other well-established areas of cognition (e.g. memory), defining and operationalising the construct of EF has been historically challenging. Numerous researchers have accepted this challenge and directed their research towards producing a definition of EF, resulting in excess of 30 published definitions of EF to exist today (Goldstein & Naglieri, 2013). These definitions are anchored to either the proposed group of cognitive skillsets that comprise EF (Anderson et al., 2011a; Miyake & Friedman, 2012; Miyake et al., 2000), its neurological underpinnings (Fuster, 2000b; Pribram, 1973; Stuss, 2011), or the nature of its behavioural response processes (Barkley, 2012; Lezak et al., 1995; Shallice, 2002).

There is growing recognition that EF is multifaceted, comprising of a number of cognitive abilities that engage interdependently to fulfil a variety of complex task demands. Furthermore, EF is no longer considered a singular construct within the frontal lobes, and is accepted to involve a diverse range of cortical and subcortical regions in its function (Alvarez & Emory, 2006; Bellebaum & Daum, 2007; Collette et al., 2006; Collette, Olivier, et al., 2005; Collette, Van der Linden, et al., 2005; Connolly et al., 2016; Lovstad et al., 2012; Niendam et al., 2012). Consequently, traditional cognitive theories that infer normality from abnormality of the frontal lobes alone have had difficulty reconciling a wide range of clinical presentations. Thus, recent discussions have advocated that singular notions of EF be retired (Logie, 2016).

To further understand this multifaceted cognitive system and how it contributes towards our engagement with complex task demands, researchers have directed their focus towards distinguishing the set cognitive skills that collectively constitute EF. This birthed a substantially large body of research focused towards the operationalization of the cognitive constructs of EF, commonly termed ‘Executive Functions’ (EFs) (Anderson, 2002; Friedman & Miyake, 2016; Hughes & Graham, 2002; Lezak et al., 2012; Miyake & Friedman, 2012; Oosterlaan et al., 2005). This approach has provided both academic and clinical utility for the creation of neuropsychological assessments that are designed to establish complex, novel and goal-orientated task environments. This has further served to assist in the identification of deficits to particular EFs across numerous psychological disorders (Barnard et al., 2008; Bott et al., 2014; Chan et al., 2009; Dixon et al., 2004; Huey et al., 2009).

A central challenge amongst the literature that has sought to define a set of EFs is the ability to achieve balance between parsimony and universal applicability of its constructs. Many models of EF have demonstrated partial success by offering *what* EFs are required during complex task demands, or *how* EFs contribute towards successful engagement with complex and novel demands. However, the two rarely feature alongside one another. These conceptual models of EF are ultimately inferred from a collection of assessment tasks that are considered to predominantly require distinct EFs for completion. This has provided an attractively linear application between theoretical EF models to clinical assessment via the notion that a singular outcome score of a task represents performance of a single Executive Function. However, many assessment tasks of EF are continuously critiqued for their inability to provide a comprehensive method of inquiry due to issues of measurement impurity, poor ecological validity and poor psychometric validity.

This very critique serves to represent the current dissonance between our growing recognition of this multifaceted cognitive system, and our continued desire to conceptualise and measure its components in isolation. The study of EFs has been successful at eluding as to what predominant Executive Function may be shared amongst a set of assessment tasks. However, this approach fails to offer context to their performance in relation to the varied demands for controlled behaviours that are required to be enacted in response to the task. Such information may be available qualitatively via observation, with deficiencies seen when a particular task demands change. However, the absence of a compelling empirical alternative has forced a reliance on traditional administration and scoring outcomes to inform both theoretical and assessment models of EF.

The notion that task performance can be considered equal or exclusive to a component of EF persists, despite neuroscientific and behavioural evidence demonstrating otherwise. However, while our ability to capture this information from current assessment tasks remains unconsidered and unavailable, this underrepresentation of human cognition is likely to continue. Given that current EF conceptualisations remain mired in past notions that one task measures a singular component, advancing cognitive theory will remain elusive until researchers cease to use unsophisticated administration and scoring to assess a poorly defined homunculus.

A related, yet distinct approach to further understand human engagement with controlled goal-directed behaviour has been offered via advancements in the field of neuroscience. Investigations into the neural architecture of complex behaviour have focused on task demands to better understand neural recruitment during tasks purported to require EF, rather than traditional administration and scoring. The emphasis within this body of research has been towards understanding neurological engagement in response to demands both *within* and *between* the execution of EF tasks, instead of a traditional reliance of between task comparisons and outcome scores only. In doing so, the problem of *what* and *how* a task is measured is diminished, and the various environmental demands from a task that can induce neurological change due to the increased need to exert control over a response come to the foreground.

As imaging methods have become more sophisticated and the body of research using these methods has grown, so too has the promise of a task demand approach. This approach has increased our understanding of how cognition is controlled and managed during complex and novel environments, that surpasses the anatomical isolation of its neural components, and acknowledges the adaptability and functional recruitment of neural regions across the cortex during many traditional EF tasks. The progressive insight that is offered by cognitive control theory highlights the misalignment between current neuroscientific evidence and the theoretical representation of EF within current cognitive theory, the diverse environment that exists within its measures, and the current gap between clinical practice and research settings. An approach that acknowledges the recruitment of neural structures across the cortex shows promise in bridging this gap, however further evaluation of its application is required.

It is the position of this thesis that the current dominancy of EF to account for complex goal-directed behaviour, its measurement, and theoretical accounts exist at a precipice. The continued utilisation and approach to the assessment of this sophisticated human behaviour must not only acknowledge the existence of a multifaceted system, but begin to reconcile the multiple lines of evidence that can serve to inform it, if the very nature of this system is to ever become demystified. Our understanding of the neurological bases of complex behaviour has evolved to recognise the importance of task demands when attempting to understand complex human cognition. It is imperative that theoretical accounts of EF embrace this understanding and evolve to reflect contemporary neurological evidence in order to improve the continuity of their conceptualisations amongst the research and clinic setting.

This thesis aims to offer a Demand Classification System (DCS) that can be applied to Executive Function assessment tasks to further understand the unique contexts and demands that call for controlled behavioural responses by EF and other cognitive systems. Many EF tasks are well suited to establishing the complex and novel task environments that require a collection of controlled behavioural responses. The DCS recognises that the most complex tasks performed by human cognition are multifaceted and aims to provide a useful framework by conceptualising performance along dual continuums of task novelty and complexity. By breaking down a task into its constituent parts using this framework, we can better understand performance across an influential hierarchy of demand that occurs independently of the specific

EF being measured. This adds to existing notions regarding the nature of EF by contextualising traditional accounts of EF with contemporary accounts of cognitive control. This serves as a preliminary step toward a unified conceptualisation that recognises the behavioural adaptivity required across varying demands of goal-directed behaviour, which may ameliorate long-standing challenges with theoretical heterogeneity, ecological validity, and measurement impurity reported amongst the EF literature. Only then can a model of EF emerge that is reflective of the core adaptability that humans endure in complex goal-directed environments.

Chapter 1

Cognitive Control and the Environment

As the modern world continues to advance and evolve, the boundaries of human cognition are continuously challenged by the intake of new information within complex and novel environments. The cognitive processes that enable humans to successfully engage with, and master, complex environmental demands are considered a central hallmark of human capabilities (Barcelo et al., 2006; DeFelipe, 2011; Pezzulo et al., 2014). Much of this engagement is carried out in order to achieve a future goal, and to plan and carry out the sequence of thoughts and actions to achieve it (Badre et al., 2010; Desrochers et al., 2019; Farooqui & Manly, 2019; Mäki-Marttunen et al., 2019). Goal-directed behaviours may be simplistic or lower-order - such as reaching into a draw to retrieve a particular item, or increasingly sophisticated (higher-order) - such as building a computer. What exists between the goal and the set of actions required to achieve it is a vast continuum of varying complex cognitive and behavioural processing that is required in order to reach a satisfactory outcome. The term 'cognitive control' has been sanctioned to describe the cognitive processes that enable, execute and manage goal-directed behaviours (Badre et al., 2010; Barcelo et al., 2006; Fuster, 2000b; Koechlin et al., 2003).

Cognitive control enables the human capacity to manage goal-directed behaviours, particularly when numerous response actions are available. This may be driven by the use of internal (endogenous) representations that guide the correct thought and action in novel situations, or conversely, may be enhanced or supported by exogenous environmental cues (Badre & Nee, 2018; Bocanegra & Hommel, 2014; Botvinick, 2008; Duncan et al., 1996; Jeon & Friederici, 2015; Miller & Cohen, 2001).

One hallmark feature of cognitive control is its *adaptive* capability to update goals and strategies when internal or environmental factors indicate that it may be advantageous to do so

(Mäki-Marttunen et al., 2019; Surrey et al., 2017). However, many everyday behaviours that are carried out via routine, habit and automatic processes can also require the activation of cognitive control mechanisms (Bargh & Morsella, 2008). Therefore, cognitive control is necessary, when internal plans and goals are required to be enforced, or by contextually deviant or novel events that present themselves (Badre & Nee, 2018; Barceló & Cooper, 2018; Barcelo et al., 2006; Duncan & Owen, 2000; Escera et al., 1998; Janowich & Cavanagh, 2019; Lavie et al., 2004; Pisula et al., 2019; Ranganath & Rainer, 2003; Waskom et al., 2014).

A significant body of neuroscience research now exists that has investigated the nature of cognitive control under various environmental demands. Key features of the environment that require cognitive control have been identified and indicate the most overarching guiding concepts to be the extent of complexity and/or novelty of the environmental demand (Stuhr et al., 2018). To understand the cognitive control processes that influence behaviour, the elements of complexity and novelty must first be defined.

1.1 Features of a Complex Environment

A central determinant to the complexity of the environment for goal-directed behaviour is the demand for adaptiveness. Within a simple environment, demands for adaptiveness are typically low due to relatively clear stimulus and response (S→R) relationships, basic instructions, or requiring little decision making amongst potential alternatives. For an environment to be complex, a challenge must be raised by the environment that requires the individual to adapt beyond direct engagement. Researchers have investigated cognitive control through the lens of individual adaptiveness across many key environmental demand features. The recruitment of cognitive control has been attributed to environmental demands for the abstraction of underlying relationships to produce an effective response (Badre & Nee, 2018; Coutlee & Huettel, 2012; Duncan, 2013; Duncan et al., 1996; Jeon, 2014; Nee et al., 2014), management of contextual changes that surround the behaviour (Aben et al., 2019; Janowich

& Cavanagh, 2019; Murphy et al., 2018; Schmidt & Lemercier, 2019; Sheu & Courtney, 2016; Siman-Tov et al., 2019), determination of an appropriate decision from many potential alterative actions available (Brown et al., 2008; Smith & Ratcliff, 2004; Yoshida & Ishii, 2006), application of necessary instructions or rules (Brass et al., 2017; Cohen-Kdoshay & Meiran, 2007, 2009; Cole et al., 2018; Hazeltine et al., 2011; Longman et al., 2019; Wenke et al., 2015), and for management of simultaneous actions (Fischer & Plessow, 2015; Olszanowski & Szostak, 2019).

Importantly, the collective presence of these complex features is not necessarily required to determine an environment as complex. This variability is arguably reflective of natural world states, where environmental demand features can be diverse, but ultimately can collectively amass to demand cognitive control resources at varying levels. Therefore, the divide between simple and complex environments can be somewhat enigmatic if demands across specific features of the environment are not considered first. To do this, these specific features must be identified in their own right in order to further understand how cognitive control is responsive to the demands that are placed upon it.

1.1.1 Abstraction

Unique to human cognition is the capacity to form both short-term and long-term abstract goals whereby abstract relationships between multiple different stimuli are formed to reach a series of short-term subordinate goals, that together ultimately fulfil a long-term superordinate goal (Cooper & Shallice, 2000; Fishbach et al., 2006; Höchli et al., 2018; Niki et al., 2019; Taylor et al., 2006). Acknowledgement of the human capacity to form abstract relationships within the environment dates as far back as the works of Plato (Saracco, 2017). Abstraction refers to the degree to which processing/representation is tied to or separated from particular instances (Nee et al., 2014). Höchli et al. (2018) suggested that one of the most fundamental characteristics of a goal is the level of abstraction that is required. The lowest form of abstraction requires a direct interaction and execution of subordinate goals that define precisely what and how to do a particular action (e.g. whisk the sugar and flour together) (Carver & Scheier, 2001), whereas superordinate goals often encompass a more complex overarching goal state, requiring overall higher levels of abstraction (e.g. bake a three-tiered cake) (Höchli et al., 2018). Using this example, the goal of making a cake serves as the superordinate long-term goal that requires numerous subordinate (sub) goals to be completed. These goals must be structured in a way that allows for both accuracy and efficiency of the controlled behaviours required. Specifically, the advent of neuroscientific technologies has brought about a set of contemporary operational definitions of abstraction requirements that comprise of several key alternate forms; temporal abstraction, policy abstraction and relational abstraction (Badre & Nee, 2018).

During *temporal abstraction*, the individual is required to simultaneously formulate and manage a sequence of subgoals with specific time-tags, and in a set order that affords the achievement of a superordinate goal(s) at a later time (Badre & Nee, 2018). When actively engaging in an environment that requires temporal abstraction, actions that can be directed by existing environmental cues are considered somewhat less demanding, in contrast to actions that require the integration from remote time-based events in the environment. To extend the analogy of the cake, the mixing of dry ingredients must be completed first, followed by wet ingredients in order to reach the desired end result. Thus, the temporal arrangement of sub-goals demands consideration of future superordinate goal attainment.

Policy abstraction refers to the evaluation of relationships between the stimulus, the appropriate action, and the consequence of taking that action. Badre & D'Esposito (2009) proposed that policy abstraction is best understood as containing lower and higher-order abstract policies. Lower-order policies often contain simple exogenous rules that link a stimulus to a response (Badre & D'Esposito, 2009). For example, when driving and seeing a green light, you may go, but when seeing a red light, you must stop. In the event that both lights are present, stimulus and response (S→R) policies are in conflict and require a higher-order policy order to determine action. The presence of context-dependent rules such as this requires the abstraction of a higher-order policy to enable the most appropriate decision to be reached (Badre & D'Esposito, 2009). In this event, the higher-order abstract policy can over-ride lower polices to go when the traffic is clear. Thus, the situation is more abstract whereby an overarching rule (e.g. traffic is clear, no pedestrians crossing, traffic lights are not operational) govern the acceptance of lower-order policies (to go, or not to go) Badre et al. (2010) found that when higher-order abstract policies were available, individuals adopted this policy, and did so rapidly.

Relational abstraction occurs when separate features of a particular dimension need to be identified in relation to one another prior to decision-making (Badre & Nee, 2018; Davis et al., 2016). Relational abstraction can vary depending on whether the properties of the stimuli are concrete (e.g. "do numbers match?"), versus drawing relationships between relationships, such as analogies (e.g. "Car is to road, as boat is to _____"). Researchers have found that increasing the number of independent dimensions or stimuli that require relationships to be considered, increases the level of abstract relation required, and manifests as additional temporal delays in decision-making processes (Kroger, 2002; Nee et al., 2014).

Many tasks often require multiple forms of abstraction in simultaneous execution for successful completion (Nee et al., 2014). As task requirements and contexts become more complex, additional contexts may need to be generalised over more rules (policy abstraction), features of the tasks may be required to be integrated (relational abstraction), and the overall goal or context must be sustained over time while lower-order decisions and sub-goals are completed (temporal abstraction) (Badre & Nee, 2018).

1.1.2 Contextual Stability

When responding to complex task environments, the individual often needs to anticipate and form expectations about what will occur next, prior to information becoming fully available within the environment (Murphy et al., 2018). The availability of contextual information is considered a key facilitator to inform associations between environmental stimuli and potential behavioural actions (Siman-Tov et al., 2019). Early research by Palmer (1975) found that when presented with an environmental context (e.g. a library) and an appropriate contextual stimulus (e.g. a book), prior expectations of intended action (e.g. reading) can be formed. This perceptual encoding affords for preparedness of a given action (Palmer, 1975). With practice, regularities in the environment can present anticipation for contextual cues that enables the learning system to associate the contextual cue with action behaviour, so that when similar cues are congruent with the environment, the behavioural response can be automatic (Berkman, 2018).

Contextual associations can be formed through habit, or learning made explicitly via instructions (Barceló & Cooper, 2018). However, without exogenous rules the individual must rely more heavily on internal representations of the context and appropriate action (Janowich & Cavanagh, 2019). Many everyday engagements are considered to favour a combination of experience with a given context, and new contextual information that must be amalgamated to form a new context representation. These may include actively represented information about prior or current stimuli, as well as more abstract information such as instructions, rules or goals, that are required to guide behaviour in accordance to current demand (Janowich & Cavanagh, 2019). Once the contextual representation and chosen behaviour are evaluated as appropriate, or retrieved from episodic memory, the individual is required to maintain the task-relevant contextual information over time until completion of temporally extended behavioural goals (Badre & Nee, 2018; Janowich & Cavanagh, 2019).

When changes occur in environmental stimuli, Surrey et al. (2017) demonstrated that controlled behavioural performance could be largely unaffected as long as the overall context remains unchanged. However, when contextual information does change, this can create task uncertainty about what is to be done with target stimuli (Barceló & Cooper, 2018). From their review of the literature, Barceló & Cooper (2018) concluded that when individuals are required to either sort a series of stimuli in one context (e.g. colour only), or between two contexts (e.g. colour and form) without exogenous cues, behavioural performance is negatively affected. This negative impact was considered to emerge due to the absence of contextual cues to provide contextual information on both temporal action (when) and task action (what) (Barceló & Cooper, 2018).

One such contextual change that is considered to require high levels of cognitive control is when a previously congruent context representation becomes incongruent with the behavioural action required. For example, the context of the library presents a quiet space where congruent behaviour would be keeping a low and quiet voice, whereas an incongruent contextual stimulus would be to hear people reading aloud. The impact of additional cognitive control due to contextual shifts often manifests as slower response times to enact a newly required action (Aben et al., 2019; Schmidt & Lemercier, 2019; Sheu & Courtney, 2016).

1.1.3 Action Rules

Despite abstract goals being formulated and contextualised, the actions that are to be taken to reach it are not always clear. When two or more alternative options are available, the individual is required to make a choice amongst action rules. The availability of the choice is governed by the presence or absence of contextual rules and explicit parameters within the environment that determine the number of perceived action rule choices available. For example, when following a map, an individual may need to choose between travelling in a left or right direction (two parameters), or continuing forwards or backwards in addition to left or right options. During these circumstances, the knowledge that is related to the correct identification of the required sequences may be accessible via external habits, schemas, plans or scripts (Dezfouli & Balleine, 2013; Norman & Shallice, 1986). However, when this is unavailable, the appropriate choice between two or more alternative responses demands evaluation, and cognitive control is required (Vohs et al., 2008).

When alternative actions are detected within the environment, several theorists (Brown et al., 2008; Smith & Ratcliff, 2004) propose that the information about each environment can be accumulated as 'noise' during the decision-making process. This noise is accumulated until evidence in favour of one particular action reaches a threshold that determines its execution (Brown et al., 2008; Smith & Ratcliff, 2004).

When the value of each task or action is not able to be determined or augmented via contextual information, response ambiguity can arise. Ambiguity is considered to represent a state of uncertainty about the probability of choosing an accurate specific action from the available alternatives (Hsu et al., 2005). The effects of environmental ambiguity are often revealed during behavioural based maze tasks. These tasks require the individual to navigate a hidden maze to reach a goal state. The hidden maze creates ambiguity in the response, with decisions needing to be based on current beliefs of the task set in addition to the incoming information from the environment stimuli (Yoshida & Ishii, 2006). Importantly, the 'noisier' the information (i.e. the more stimulus input, alternatives, uncertainty and ambiguity) the longer decision thresholds take to reach, which corresponds with a trend of longer reaction times and reduced accuracy in the execution of behavioural action (Smith & Ratcliff, 2004).

1.1.4 Instructions and Rules

Human beings are able to assimilate novel information about tasks and action behaviour from mere instruction and transform it into symbolic goal-directed behaviour (Cole et al., 2011; Duncan et al., 1996; Liefooghe & Verbruggen, 2019; Verbruggen et al., 2018). Explicit instructions help to prepare both motor and perceptual systems to facilitate successful actions (Cole et al., 2013; Liefooghe et al., 2012), and avoid less effective and slow responding to environmental requirements (Cohen-Kdoshay & Meiran, 2007, 2009; Longman et al., 2019; Palenciano et al., 2019). Essentially, instructions facilitate automaticity and therefore the availability of instructions may not elicit much cognitive control in order for task completion (Brass et al., 2017; Wenke et al., 2015).

According to Liefooghe & Verbruggen (2019), instructions enable the establishment of a task model set. The cognitive processing required for the establishment of this model set can be minimal when the construction and preparation of S→R actions are informed directly from instructions (Brass, 2002; Brass et al., 2017; Hartstra et al., 2012; Ruge & Wolfensteller, 2010). Thus, when instructions of how to proceed with a sequence of actions are available, the demands of abstraction, contextual processing and action rule determination may be substantially decreased (Sudevan & Taylor, 1987).

Several authors have highlighted that the provision of rules is not deterministic of outcome or ease of the task, as instructions that are required to be memorized place sufficient load on cognitive reserves that will demand cognitive control (Brass et al., 2017; Meiran et al., 2015; Muhle-Karbe et al., 2016; Paas & Van Merriënboer, 1994). Muhle-Karbe et al. (2016) established two task settings, one in which action responses were directed by preceding instruction across primary and secondary tasks, and a second setting when instructions changed across tasks. It was found that instructed stimuli gained the power to automatically activate the associated responses in a secondary task, only when the instruction was prepared, but not when it was memorized. Within their experiment, Muhle-Karbe et al. (2016) suggested that the instructions available reduced the number of available perceived action rules, which caused the instruction to form part of the overall action plan. This is in contrast to less direct instructions, which call for the instruction set to be maintained in memory until it can later be matched and applied. Similarly, the works of Altmann (2004) and Altmann & Gray (2002) demonstrated that when instructions are given that cue sequential actions are required to be maintained in memory, these cues can often decay, exerting a negative impact on overall performance.

Importantly, when instructions are successfully implemented, any inherent instruction-based rules are then often required to be maintained throughout the task until completion of the overall established task goal (Verbruggen et al., 2014). The provision of instruction-based rules serve to confine the availability of any perceived relevant stimulus features, which can further contribute towards the establishment of action rules for each subordinate goal. For example, 'Move X prior to Y before proceeding to Z'. Verbruggen et al. (2014) proposed that a clear distinction exists between a task goal and a task rule, whereby task goals serve to describe what one sets out to achieve, and task rules help to specify how one can achieve it. Thus, rule dependent environments can induce the need for proactive cognitive control to substantiate the correct S→R relationships. (Cole et al., 2018; Hazeltine et al., 2011). When instructions and rules are absent from any environmental cue, leaving any explicit direction for available action rules unclear or uncertain, the appropriate behavioural response must be extrapolated from abstract relationships (e.g. Policy Abstraction) (Badre et al., 2010).

1.1.5 Dual-Task Requirements

The environmental conditions discussed thus far have primarily fallen within the context of completing one set of S→R actions towards a task goal. However, human interaction with the environment frequently requires that more than one task be completed at the same time. During this multitasking event, cognitive control is required to monitor behavioural actions and the demands required from each task individually in order to minimise the interference of each task on the other (Fischer & Plessow, 2015). To efficiently multitask, Fischer & Plessow (2015) proposed that cognitive control enables flexible adjustment to the environment where different processing strategies are required by each task. The increase in these control requirements often results in severe performance costs, which can manifest as reduced accuracy, and an increase in the overall delay of the responses (Fischer & Plessow, 2015; Pashler & Johnston, 1989). Importantly, while some tasks when performed individually may not typically require high levels of cognitive control (e.g. a singular exogenously instructed S→R task), when performed concurrently with another task of similar demand, performance is found to decrease dramatically due to a dual-task cost (Hommel, 1998).

Recently, Olszanowski & Szostak (2019) found that when two tasks (primary and secondary) have a limited number of action rules available, performance accuracy was mostly unaffected. This was attributed to the task environment affording the automation of task responding by allowing for a reduction in decision-making requirements of both tasks, thus reducing dual-task costs. However, when increasing individual cognitive control demands in the primary task by reducing automation responding, a decrease in performance accuracy is reported across both the primary and secondary tasks (Olszanowski & Szostak, 2019).

1.2 Features of a Novel Environment

In addition to the influence of complexity, a second key environmental feature requiring cognitive control involves the management of novelty. Human engagement with complex environments often encompasses aspects of novelty. To ensure that experiences are not redundant, heuristics and schemas are formed that are later accessed when environmental similarity is experienced. A heuristic is a mental shortcut, allowing problem-solving and quick judgments to be made efficiently. Heuristics shorten the decision-making process and decrease the level of cognitive resources required to determine the following course of action (Cole et al., 2011). In the context of controlled behaviour, schemas provide foundational knowledge that provides a mental shortcut to action rules required to adequately respond to the current scenario. Schemas allow for the transference of knowledge to new task contexts, whereby a novel task can be interpreted from a combination of familiar representations rather than having to learn each complex rule and action from the beginning each time (Cole et al., 2011). When in familiar environments, cognitive control has the capacity to be proactive, given that familiar contexts, cues and action rules may be detected that enable preparation for the anticipated S→R (Verbruggen et al., 2014). In contrast, in a truly novel environment that bares unfamiliar S→R features, cognitive control can be considered largely reactive in its activation (Verbruggen et al., 2014). It was recently proposed that the presence of instructions enables a greater capacity for proactive cognitive control when completing novel tasks (Cole et al., 2018).

The onset of such novel demands within a familiar environmental context can also cause cognitive resources to be disrupted from any ongoing task performance (Barcelo et al., 2006). Thus, when a novel situation arises due to an increase in schema acquisition required, the demand for cognitive control can increase (Wirzberger et al., 2018). This novelty may be globally influenced by a) automaticity b) schematic demands and c) episodic demands.

1.2.1 Automaticity

When a task or environmental demand is repeatedly experienced and the behaviour executed, automatization of responding is thought to develop due to consistent mapping of the S→R (Dickinson, 1985; Schneider & Chein, 2003). The enactment of automatic responses hallmarks a transition from cognitive control during an initial complex and novel task environment, to cognitive processing that is automatic and supported by long-term memory retrieval processes (Schneider & Logan, 2006; Servant et al., 2018). The automaticity of cognitive processing brought about by familiarity allows for reduced effort for interaction with the demands of the environment due to a reduction in cognitive control effort (Zink et al., 2018).

The cognitive processing of automatic S→R interactions are considered fast and parallel, effortless, and robust to dual-task interference (Bargh, 1994; Moors & De Houwer, 2006; Schneider et al., 2006). Moreover, automatic behaviours often do not require constant controlled guidance or monitoring, and therefore only minimally tax attentional resources

(Wheatley & Wegner, 2001). When the automaticity of task-specific behaviour is available, Bargh (1994) suggests that an individual may be unaware of the mental process that is occurring, may not be involved with the initiation of these mental process, and may not have the ability to stop or alter a process after initiation. Thus, in most circumstances, automatic mental processes promote efficiency in actioned behaviour.

Fundamentally, automatic behaviours encompass processes and action behaviours that reflect the success of prior learning and exposure. For example, reading, colour naming and counting require a high degree of cognitive control during the skill acquisition process (Christensen et al., 2019; Uleman & Bargh, 1989). Eventually, these abilities become well-refined skill-sets that are automatic in execution with little cognitive control or conscious thought required (Reichle et al., 2006). However, when a task environment requires an individual to counter well-learned automatic S→R behaviour due to an opposing instructed or rule-governed action, higher levels of cognitive control are required. This is often demonstrated by the popular Stroop Test, whereby the automaticity of reading is required to be suppressed in favour of a more controlled colour naming process (Bub et al., 2006; Lezak et al., 2012). Although colour naming of itself would also be considered automatic, the context of the presentation of the stimulus being in word form, including incongruity between the written word and the ink colour to be named, presents challenges to the automaticity of both tasks. Thus, the demand of novel requirements by the Stroop Test environment may encompass the application of well-learned skills sets in a novel setting, requiring the recruitment of cognitive control for adaptation into a series of controlled context-specific behaviours which collectively contribute towards new schemata for the current S→R action.

1.2.2 Schematic Demands

Schemas are superordinate declarative knowledge of lower-order features that are abstracted over multiple experiences (Badre & Nee, 2018; Bartlett & Burt, 1933). Schemas are formed from the retention of fundamental knowledge gained from experience with an environment (Ghosh & Gilboa, 2014). When environmental demands allow for a continuous direct interaction with this declarative knowledge, schemas can be accessed, and the appropriate response formulated with minimal cognitive control (Mazzone, 2015; Wirzberger et al., 2018). In contrast, novel and complex environments require increased cognitive control due to the demand for abstraction of lower-order relationships, which are then applied to the novel environment stimuli that is present (Badre & Nee, 2018).

Cole et al. (2011) suggested that during novel environmental demands, compositional schemas of rule representations can be established whereby new task representations can be constructed from different combinations of familiar rule representations (Cole et al., 2011). This is considered particularly applicable to decision-based rules which are highly abstract and compositional, and usable across a wide variety of contexts (Cole et al., 2011). Although both concepts seem similar on the surface, schemas differ from automaticity in the number of behavioural patterns that are combined whereby automaticity occurs in response to singular concepts or behaviours, and schemas are called into play in more complex or multi-layered environments.

1.2.3 Episodic Demands

During complex tasks where multiple sequences of subordinate goals must be achieved to reach an overarching superordinate goal, the retention of relevant contexts, abstract relationships, rules and instructions may be required to be retained (Veronica Mäki-Marttunen et al., 2019). The completion of one sub-goal may be instrumental in cueing the next sequence of behaviour that is planned (Badre & Nee, 2018). When task progression is dependent on retention of previous behavioural responses, episodic control may be required whereby certain ‘episodes’ of a sequence of S→R actions indicates the next sequence of behaviour (Badre & Nee, 2018). Thus, episodic control assists with the monitoring and activation of temporal demands across task sets. For example, a flashing light may indicate that the next sequence of behaviour is to either continue or stop the current behavioural action. Additionally, the environment may demand that the individual reference a point of the task that has previously occurred in order to respond to a current demand. (Badre & Nee, 2018). For example, take the answers for Task A, and apply them together with Task B.

When establishing a new task set, it is optimal to reuse previous S→R knowledge wherever possible (Domenech & Koechlin, 2015). A rapid adaptation to novel tasks can rely on abstract knowledge from prior experience with similar tasks (Bhandari & Badre, 2018; Domenech & Koechlin, 2015). With sufficient representations available, the amount of additional learning required is minimised, which can facilitate improved performance in novel circumstances (Cole et al., 2011; Cole et al., 2013). However, remembering and keeping track of multiple relevant contexts may influence the temporal dynamics of context updating, and thus increase the need for overall cognitive control as the number of sets of S→R interactions increases (Mäki-Marttunen et al., 2019). The application of instructional rule-based actions in the face of novel and changing environmental contexts demands the re-assessment of the abstract relationships between the novel S→R, together with the integration of any relevant episodic information (Badre et al., 2010; Domenech & Koechlin, 2015).

1.3 Neuroanatomy Supporting Cognitive Control

With the advancement of neuroscientific technologies, considerable progress has been made towards understanding the neural networks that support cognitive control and complex behaviour across the brain, and how these networks are engaged during various task demands. The outcome of these efforts is the identification of generalised task networks within the brain that seemingly support a range of complex behaviours (Cole et al., 2011; Duncan, 2013; Duncan & Owen, 2000; Niendam et al., 2012). For example, Duncan (2013) discovered that stable and consistent activations occur across the frontal and parietal regions of the cortex during an array of tasks that were considered to require differing cognitive demands. This observation lead to the proposal of a Multiple Demand System (MD) that is responsive to fluctuations of attentional episodes across time (Duncan, 2013). Attentional episodes refer to sub-goals that require activity across the brain to be configured for a solution to the problem at hand. The cortical regions within the MD system include the Dorsolateral Prefrontal Cortex (DLPFC), the insula cortex, precentral gyrus, anterior Supplementary Motor Area (preSMA), and the anterior and middle cingulate (Duncan, 2013). Activations within these regions are considered to vary between each individual; however, their recruitment during various conditions of goal-directed tasks are found to be fairly consistent (Fedorenko et al., 2013).

A more recent extension of the MD system is the Extrinsic Mode Network (EMN) (Hugdahl et al., 2015). The EMN is considered to be a generalised task network that is complementary to the Default Mode Network (DMN). The DMN is an intrinsic control network (ICN) that encompasses a group of cortical networks that are found to be active during rest when no specific task requirements exist (Raichle et al., 2001). During a resting state, DMN activity is located within temporal, posterior cingulate, lateral inferior parietal and precuneus regions (Raichle, 2015). In the presence of increasing task demands the DMN deactivates and EMN activity becomes upregulated across rostral neural networks, including the lateral frontal and parietal regions. Activation of the EMN is found to respond consistently to tasks measuring a variety cognitive control demands that require various cognitive sub-processes for completion (e.g. working memory, inhibition, mental rotation, updating, cognitive flexibility, mental arithmetic, theory of mind) (Hugdahl et al., 2015). Therefore, EMN recruitment is considered to highlight the common shared demands that exist between complex task demands (Hugdahl et al., 2015). Hugdahl et al. (2015) acknowledge that peak amplitude will vary amongst the brain structures of this network according to different tasks demands, but the overall architecture of the EMN should remain by and large the same.

The identification of a neural architecture that remains stable during the execution of a diverse range of complex tasks offers an intriguing insight into the cortical networks shared by task demands that, at the behavioural level, appear to manifest as different aspects of cognitive control and skillsets. While this evidence is emerging and further exploration is required to demonstrate the extent to which these task-general neural systems can support specific task demands, the upregulation of these networks during a variety of task demands provides a useful insight into how regions within these networks may respond to support engagement with complex environmental demands.

1.3.1 Rostro-Caudal Organisation

The recruitment of the lateral pre-frontal cortex (LPFC) and parietal regions in response to increasing task demand complexity demonstrated by the MD and EMN is consistent with leading neuroscientific evidence that attributes the functional architecture of the cortex to be predominantly hierarchical in its ability to engage cognitive control (Fuster, 2000b; Fuster, 2017; Koechlin, 2016; Koechlin & Summerfield, 2007). Within this hierarchy, posterior regions of the cortex activate in response to sensory and orientation integration requirements, whereas frontal regions of the cortex become engaged when abstraction and cognitive control are required (Yoshida et al., 2010). The gradient of this rostro-caudal organization is found to be dependent on the automaticity of the task at hand, with frontal regions exerting activation dominancy over posterior cortical regions when novel, unfamiliar demands are present (Jeon & Friederici, 2015). A central assumption to the rostro-caudal organisation of the brain is that the frontal lobes are central to the selection and execution of controlled action, and to some extent exert a top-down control over posterior cortical and subcortical regions (Fuster, 2000b, 2001, 2017). In addition to their hierarchy within the cortex, recent evidence suggests that the organization and processing of cognitive control mechanisms within the frontal lobes to be hierarchical as well.

Koechlin & Summerfield (2007) proposed a cascade model of 'executive' control that fractionates regulation by the Pre-frontal Cortex (PFC) into hierarchical levels of cognitive signals that are activated by sensorimotor, contextual and episodic information (Koechlin & Summerfield, 2007). The processing of each signal is found to coincide with regions hierarchically along the rostro-caudal axis of the LPFC. Sensorimotor regions (BA6) are engaged in selecting a behavioural response based on prior exposure, contextual control regions (BA 9/ 44/ 45) select the appropriate sensory-motor representation in the premotor cortex in accordance with the immediate context, and episodic regions (BA 46) select an action based upon representation according to the temporal context (Minamoto et al., 2015).

Badre & D'Esposito (2007) also support the existence of hierarchical processing of increasingly complex and abstract neural representations within the LPFC. However, it is proposed that the hierarchy of the LPFC is better applied to the selection of competing representations during action selection, instead of a control signal (Badre & D'Esposito, 2007). Within this perspective, the rostral LPFC holds the abstract overarching goal that does not specify the actions required for completion. The actions required to achieve the goal are broken down into sub-goals at multiple levels of the caudal LPFC. The levels of this representational hierarchy are differentiated by the level of abstract representation (or complexity of the action rule) that must be selected over competition, with the rostro-caudal gradient emerging as the level of abstraction is increased (Badre & D'Esposito, 2007).

There is consensus amongst researchers that an increase in any abstract rule results in an increase to the hierarchical activation pattern within the LPFC (Badre & D'Esposito, 2007; Botvinick, 2008; Farooqui et al., 2012; Jeon, 2014; Koechlin & Summerfield, 2007; Minamoto et al., 2015; Nee & D'Esposito, 2016; Schneider & Logan, 2006; Yoshida et al., 2010). However, details on whether differences exist in the hierarchical gradient amongst subregions of the LPFC between different abstraction forms is still inconclusive (Bahlmann et al., 2015; Nee &

D'Esposito, 2016; Nee et al., 2014), and is likely to be attributed to the unique abstraction requirements required within a given task environment (Badre & Nee, 2018).

Recently, Bhandari & Badre (2018) discussed the importance of Working Memory (WM) control policies as a key process to support flexible behaviour in novel circumstances. Within these control policies, the contents of WM are selectively updated via an input gate that determines whether the current information can enter WM stores. A selective output gate then allows WM to influence downstream processing or the task environment. The DLPFC is considered central to enable the flexible updating and encoding of the context of controlled behaviour to WM (Badre & Nee, 2018).

Badre & Nee (2018) recently summated the body of evidence from rostro-caudal activation responses to abstraction, schematic, and contextual requirements and proposed that three distinct functional networks exist within the frontal cortex. The most caudal subdivision, the pre-motor cortex, is involved in the processing and execution of sensory-motor control. The mid-lateral PFC is then central to the enactment of contextual control, and the Rostrolateral PFC (RLPFC) exerts dominance over schematic control. The RLPFC becomes engaged when control depends upon an episodic or temporal structural context, integration and inference over multiple features, and when tracking hypothetical strategies and goals (Badre & Nee, 2018). It is proposed that the RLPFC is important in processing and transmitting this information to the control hierarchy, where the mid-DLPFC is considered to be at the apex (Badre & Nee, 2018).

1.3.2 Rostro-Caudal Functional Recruitment

Investigation of the rostro-caudal hierarchy relies upon researchers to establish task-specific environments that provide a graded increase in the demand from cognitive control. Here, external behavioural tasks are sourced that aim to control a stimulus, and neural activation is recorded while the response action is carried out. To further understand how the neural networks engage cognition, key information is required to be inferred by observing the controlled behaviour that is enacted (Fetsch, 2016). Researchers commonly employ task paradigms that aim to explore both the exogenous and endogenous controls of cognition via manipulating distinct task procedures (Barcelo et al., 2006). Collecting behavioural data on the accuracy and efficiency of a response to the set demands serves to provide further information on the effects of task demands. For example, support for the importance of rostro-caudal functional organization has been demonstrated in the context of working memory capacity (WMC). (Minamoto et al., 2015) discovered that individual differences in WMC were underpinned by the rostro-caudal relationship within the PFC. When episodic control was required (selecting an appropriate task set according to temporal context), individuals with higher WMC recorded greater activation in the DLPFC in comparison to those with low WMC (Minamoto et al., 2015). Higher WMC also coincided with increased activation of the right inferior parietal cortex and right middle/inferior temporal cortex which demonstrates not only the extent to which the PFC exerts control during WM demands, but also the support that is required from the involvement of posterior neural structures (Minamoto et al., 2015).

When considered together, the rostro-caudal functional organization within the PFC and the progressive upregulation of this region during increased task demands can be usefully applied when interpreting comparative task-specific demands during behavioural measures of cognitive control. For example, Connolly et al. (2016) explored the ICNs involved in a common measure of EF, the Stroop Test. As previously explained, The Stroop Test is amongst the most popular interference tests that employs a congruency/incongruency switching paradigm. During the task, abrupt changes require the suppression of an automatic behavioural response in favour of a controlled and incongruent condition. During the Stroop Task, activation patterns were recorded in frontoparietal control networks, dorsal attention, sensorimotor, and visual networks (Connolly et al., 2016). Activation of the DLPFC, inferior frontal gyrus and anterior cingulate gyrus demonstrated a trend of sensitivity to the Stroop effect over the other ICN regions identified (Connolly et al., 2016). As the Stroop Effect encompasses the most complex component of the task, the recruitment of the rostral PFC regions is unsurprising given the evidence demonstrating its progressive recruitment when requirements of cognitive control are beyond the level of automaticity. At the behavioural level, incongruence within the Stroop Effect trial increases the complexity of the task due to the need to inhibit a prepotent response over the well-practised process of reading. The recruitment of the LPFC and posterior ICN's during the Stoop Task also highlights the neural networks specific to the demands of the task, and the levels of complexity that can be measured within the different conditions of the paradigm itself.

When considering the activity of ICN regions and the upregulation of a rostral-caudal gradient across the cortex during different stages of tasks, the shared and unique cognitive demands placed on the control systems can be further brought to light. Taking this approach provides clarity to how different tasks and their conditions vary in the cognitive demands they require. For example, comparing 'easy' vs. 'hard' conditions within tasks purported to measure EFs of visual and verbal working memory, inhibition, cognitive flexibility, and interference control tasks demonstrates that greater demands during the 'hard' condition results in increased bilateral activity within the precentral gyrus, middle frontal gyrus, inferior and superior parietal cortices, insula, and the opercular part of the inferior frontal gyrus (Fedorenko et al., 2013). While the 'harder' condition of each task varied depending on the task paradigm, it ultimately required an increase in the retention of information or available actions. Overall, the neural regions activated during the harder conditions provide supporting evidence for the existence of the MD system (Fedorenko et al., 2013). In addition, the increased neural activation in response to 'hard' task conditions represents how this network responds to increased cognitive demand.

1.4 Cognitive Control and Executive Functioning

The term ‘cognitive control’ has historical origins amongst early theoretical accounts of frontal lobe pathologies (Posner & Snyder, 2004). However, the foundation of modern-day cognitive control research also has close connections to the field of EF. The application of behavioural assessment tasks purported as longstanding measures of EF is often implemented to explore facets of cognitive control, which has served to demonstrate a theoretical overlap between these two neurological and neuropsychological constructs. Behavioural assessment paradigms of EF have a long history of inferring efficiency, accuracy and performance of higher-order frontal lobe derived cognition. These traditional assessment tasks have provided a useful tool towards the neurological investigation of cognitive control due to the varying degrees of controlled behaviour that they frequently demand.

While neurological cognitive control studies often utilise traditional neuropsychological tasks to infer the neural underpinnings of endogenous cognitive control, in subtle contrast, EF research utilises these same tasks to infer performance of a particular Executive Function (or Executive Skill), such as ‘planning’ or ‘inhibition’. The introduction of neuroscientific technologies has demonstrated that performance on many EF assessment tasks is underpinned by the same networks also subsumed by cognitive control including the left PFC, DLPFC, and RLPFC (Jurado & Rosselli, 2007). Arguably, the increase in both the availability and technical capacity of functional neuroscientific imaging has enabled the investigation of the activity of these neural networks during complex and novel environmental demands beyond anything to date. This has bypassed the approach of the traditional EF paradigm to uncover both unique and shared networks that respond to complex novel demands. Consequently, the popularity and return of the terminology of ‘cognitive control’ in conjunction with complex goal-directed demands is observed.

The significant increase in cognitive control research has seen the terms ‘Executive Functioning’ (EF) and ‘cognitive control’, or the amalgamated ‘Executive Control’, used interchangeably throughout the literature (Nyongesa et al., 2019). While both Cognitive Control and EF may not be mutually exclusive, they nonetheless demonstrate alternate approaches for understanding how humans engage with complex, novel and goal-directed environments. Historically, EF has been of central focus given the clinical need to identify, understand and measure behavioural pathologies associated with dysfunctions of higher-order cognition. This history is rich with the development and implementation of a multitude of tests subscribing to various paradigms and theoretical models of EF. More recently, cognitive control research is pioneering a contemporary way to understand goal-directed and controlled behaviour at the neurological level, and its connection, application and measurement in the realm of EF requires further understanding so that such advancements offer utility to the cognitive and clinical sciences.

Chapter 2

Executive Functioning & Controlled Behaviour

The offering of a theoretical account of EF that provides a parsimonious yet comprehensive account of the nature of higher-order cognition has remained largely elusive and heavily contended, despite exercising the literature for a number of decades. Nonetheless, there is general agreement that the nature of EF facilitates the control of goal-directed behaviour (Anderson et al., 2011; Barkley, 2012; Best & Miller, 2010; Dawson & Guare, 2010; Lezak et al., 2012; Luria, 1976; Miller & Cohen, 2001; Oosterlaan et al., 2005; Welsh & Pennington, 1988) to respond adaptively to novel (Barkley, 2012; Banich, 2009; Chan et al., 2008; Gioia et al., 2000; Norman & Shallice, 1986; Snyder et al., 2015) and complex environment demands (Diamond, 2013; García-Madruga et al., 2016; Hughes & Graham, 2002; Jurado & Rosselli, 2007; Lezak et al., 2012).

Unlike many other well-established areas of cognition (e.g. language and memory), the functional processes of EF and its facets have historically been difficult to isolate at the neurological level. Consequently, their representation in cognitive models has been heterogeneous and heavily debated. After all, the success of a cognitive model is demonstrated by its ability to account for the behavioural consequences of neurological dysfunction, which for EF has proven a considerable challenge. A long-standing awareness has existed regarding the involvement of the frontal lobes in EF and higher cognition (Anderson et al., 2011; Fuster, 2017; Luria, 1976; Stuss, 2011), with early observations of frontal lobe pathologies which implicated their involvement in achieving goal maintenance, planning, and coordination of complex behavioural sequences (Luria, 1976; Stuss & Benson, 1984).

The existence of over 30 published definitions of EF demonstrates the continuing lack of consensus towards how EF is best conceptualized (Goldstein & Naglieri, 2013). The ongoing debate surrounds *what* cognitive skills/subcomponents, termed 'Executive Functions' (EFs), are subsumed within the overarching construct of EF. These functions are commonly considered to be required for, or a product of EF, that are measurable using neuropsychological tests. Varying models sometimes depict similar, but more often than not opposing positions on what specific functions accurately reflect EF as a macro-construct. Nonetheless, leading models postulate that these are the abilities that are called into service when an individual is required to complete a novel and/or complex task (Shallice, 2002).

2.1 Goal-attainment Models and Research

Many traditional theories of cognitive control & complex goal-directed behaviour (now commonly referred to as theories of EF) were established from inferring normality from abnormality of the frontal lobes (Barkley, 1997). This approach resulted in the offering of cognitive models that focused on a limited range of behavioural processes. Early observations of frontal lobe pathologies influenced the production of cognitive models that inferred the situations where effortful control is required (Goldstein & Naglieri, 2013). Posner & Snyder (1975) (As cited in Posner & Snyder, 2004) were amongst the first researchers to propose that the mechanism of cognitive control was responsible for overwriting automatic responses. Within this model, cognitive control allows an individual to adapt to a situation depending upon their intended overall goal. Shiffrin & Schneider (1977) further extended this understanding and proposed that control processes are needed in order to attend to and select certain stimuli in favour of others. This seminal research provided the early foundations for our current understanding of automatic vs. controlled processing.

Subsequent to these conceptualizations of cognitive control processing, researchers called for an understanding of what enabled the selection and coordination of controlled behaviour. Baddeley & Hitch, (1974) and Baddeley (1996) proposed a solution that propositioned the existence of a Central Executive (CE) that operated within the frontal lobes. Similar to other early models of cognition, a limitation of this model was the attribution of a unitary controller, often described as a 'Homunculus', which regulated lower-level systems. Subsequently, in light of the diverse EF symptomology observed from various neurological pathologies and the inability to locate a single neural substrate of EF, the CE approach has been unable to account for the varied clinical presentations that are observed with EF dysfunction. While this approach paved the way for future understandings of 'what' is controlled, such accounts did not depict 'how' it is controlled (Verbruggen et al., 2014). Therefore, it was more recently suggested that the notion of a CE be finally retired (Logie, 2016).

2.1.1 Process Driven Models

The need to further understand how EF is controlled has resulted in the fractionation of the CE concept in an aim to understand how distinct processes underlie and regulate controlled behaviour (Verbruggen et al., 2014). The Supervisory Attention System (SAS) proposed by Norman & Shallice (1986) provided an influential framework for acknowledging *how* differences between controlled and automatic processes can be distinctly conceptualised. This model acknowledges the construction of schemata (schemas) in novel situations to problem solve goal-directed behaviours. These schemata are then required to be implemented, which demands ongoing monitoring. If considered ineffective, the schemata are then required to be either rejected or modified. The SAS perspective holds to the notion of a central capacity within the frontal lobes that has been attributed to the role of the CE (Baddeley, 1996), and as such has suffered similar limitations Baddeley's (1996) CE approach. In an effort to overcome this, Stuss et al. (1995) proposed that five frontal supervisory processes exist that select the appropriate behaviour that is required when faced with a novel situation. Broadly, these included (1) energization of target schemata, (2) inhibition of inappropriate schemata, (3) adjustment of contention scheduling, (4) monitoring of the schemata, and (5) if-then logic to analyse monitored feedback to either maintain or alter schemata.

To further account for how goal-directed behaviour is controlled, Fuster (2000) proposed a model of cross-temporal synthesis that was based upon three concepts: interference control, planning, and working memory. Fuster's (2000) theory proposed that the primary goal of EF lies within organizing behaviour. Contrasting from previous models, such as the Baddeley (1996) WM model, Fuster (2000) does not *place a ghost in the machine* (Barkley, 2012). There is no CE or singular component within the model of cross-temporal synthesis; rather temporal mediation captures the interaction between short-term memory and the attention set (Fuster, 2000). However, a common limitation to these early process models was the further need to understand the components of EF that could be observed and measured during clinical assessment. This need impelled researchers to actively focus on identifying the collection of measurable cognitive domains that are utilized during complex, goal-directed tasks.

2.1.2 Domain Centred Models

The understanding that EF is required for successful completion of goal-directed behaviour in complex & novel environmental demands gave rise to a substantial body of research focused towards operationalising the cognitive constructs and process of EF, the EFs (Anderson, 2002; Friedman & Miyake, 2016; Hughes & Graham, 2002; Lezak et al., 2012; Miyake & Friedman, 2012; Oosterlaan et al., 2005). This approach has provided both academic and clinical utility for the creation of neuropsychological assessments that are designed to measure a particular facet of EF. This has further served to assist in the identification of EF deficits across numerous psychological disorders (Bora et al.; Chan et al., 2009; Elliott, 2003; Hill, 2004; Huey et al., 2009; Oosterlaan et al., 2005; Raffard et al., 2009). However, a vast heterogeneity exists between theoretical accounts in relation to what cognitive ability can be termed an 'Executive Function' and be subsumed under the conceptual umbrella term of EF.

From the consolidation of clinic assessment and behavioural observations at the time, Lezak et al. (1995) proposed a conceptual model of EF that comprised four key domains. The model placed these domains in the context of temporal order or stages of cognitive events and behavioural actions that are required for goal-directed behaviour. Firstly, (1) *Volition* requires the conscious decision to perform an action, or an intention to carry out goal-directed or future-orientated behaviour. (2) *Planning* is then required to identify a sequence of steps necessary to solve a problem or accomplish a goal end state. (3) *Purposive action* is then required for the initiation and maintenance of the steps involved in the plan, as well as the capacity to modify or discontinue the planned actions as needed. Finally, (4) *Effective performance* is required to continuously monitor, self-correct, and regulate the action. This model has proven useful for helping clinicians to broadly identify EF components, however it suffers from a lack of specificity as to what may contribute towards each key process. For example, working memory was demonstrated earlier (Baddeley & Hitch, 1974) for being important to the enactment of controlled behaviour. Lezak et al. (1995) understate WM in their framework, whereby it would be required during all stages (Anderson et al., 2011). Impulse control, which is also found to be central to controlled behaviour, is also inherent in planning, but not explicitly stated or discussed within the model.

Anderson (2002) aimed to achieve a new model that ameliorated problems with previous accounts by proposing a conceptualisation of EF as an overall control system that is comprised of four broad domains which house a subset of cognitive skills. Anderson (2002) used the developmental trajectory of the frontal lobes as a framework to propose the domains of Attentional Control, Cognitive Flexibility, Goal Setting, and Information Processing. These four domains are separable, given that their maturation occurs at different stages of development, with an acknowledgment that they interact to a certain degree depending on the task at hand (Anderson, 2002). Attentional Control is considered to be a fundamental skill that exerts unidirectional influence upon all three other domains. Goal Setting, Cognitive Flexibility and Information Processing are then conceptualised to share bi-directional influence upon each other (Anderson et al., 2011). This representation of EF indirectly infers a hierarchical structure to EF as the involvement of attention control is considered a foundational requirement for the successful execution of the other three EF domains. The bi-directionality of the relationships between the remaining domains also infers that there is some overlap or unity of skills required between EFs.

Seminal research by Miyake et al. (2000) also confirmed the existence of separable EFs by applying confirmatory factor analyses to a range of frontal lobe tasks that were considered to elicit the EFs of Updating, Inhibition and Shifting. These domains were found to be statistically independent, although the shared variance between them was considered to contribute substantially to total EF variance (Miyake et al., 2000). This shared variance was considered indicative of the non-EF processes that are common between each of the domains that were investigated. It was suggested that these processes were associated with that specific task context (e.g., colour processing, articulation speed) (Miyake et al., 2000). As a result, Miyake et al. (2000) conceptualised a 'Unity and Diversity' framework of EF whereby unity refers to variance accounted for by the separable domains, and diversity refers to the unique variance of each.

The unity/diversity perspective has received support for the existence of separable EF factors, but with some studies finding additional factors outside of those identified by Miyake et al. (2000) (e.g. Fisk & Sharp, 2004; Fournier-Vicente et al., 2008). Further investigation by (Miyake & Friedman, 2012) revealed that the Inhibition factor within their model became non-significant in the presence of a Common EF factor that predicted task measures of inhibition, updating and switching. Common EF is hypothesized to encompass facets of inhibition and goal maintenance that are often required across many EF tasks (Friedman & Miyake, 2016).

Friedman & Miyake (2016) acknowledged that the EF factors identified within their unity/diversity model do not provide a comprehensive account of all EF domains that are assumed to exist. For example, planning - which is often considered an Executive Function, represents a higher-order construct, with Updating, Shifting & the Common EF factor operating in a collaborative manner to explain performance on planning related measures (Miyake & Friedman, 2012). However, an empirical account of the extent of this engagement remains largely elusive.

2.2 Identifying Executive Functions

It is apparent that there are distinct parallels between cognitive control theory and EF models, insofar as explaining goal-attainment and complex behaviour. However, while demand features offer a detailed description of the mechanisms of cognitive control, EFs fail to explain a mechanism of EF.

While the theoretical accounts of EF presented thus far are not exhaustive of all that exists, they are amongst the most widely accepted and utilised models within the research literature. In addition to these models there is a substantial body of work documenting numerous other specific EFs beyond those named here. The most recent systematic review concerning this topic found that among 60 of the most frequently cited studies, upwards of 68 different EFs were reported (Packwood et al., 2011). By applying a latent semantic and hierarchical cluster analysis, Packwood et al. (2011) successfully reduced this body of EFs into 18 skill types. This highlights the substantial conceptual overlap between similar task-specific behaviours that have been assigned alternative semantic labels. Reasons for the vast number of EFs available within the literature for summation by Packwood et al. (2011) echoes the effects of the longstanding practice by researchers of defining EFs using a task-focused approach (Jurado & Rosselli, 2007).

Table 1 highlights a selection of EFs that are amongst the most frequently cited within contemporary assessment models.

Table 1

Descriptions of Frequently Cited Executive Functions

Executive Function	Description
Working Memory	Holding information in mind and mentally manipulating it (Baddeley, 1998)
Planning	Identifying, organising, and sequencing multiple elements required to achieve a goal (Gross & Grossman, 2010)
Inhibition	Selectively attending and focusing on what is required while simultaneously suppressing attention to other stimuli. (Diamond, 2013)
Cognitive flexibility	Shifting attention between different response sets and processing multiple sources of information concurrently (Anderson, 2002)
Updating	Constant monitoring and rapid addition/deletion of working-memory contents (Miyake et al., 2000)
Self-monitoring	Self-observation of one's own performance including measuring it against some standard of what is needed or expected (Burke et al., 2009; Wilde & Garvin, 2007)

One of the key problems plaguing EF research is that successful recruitment and performance of a particular function is measured and attributed in the form of an often singular outcome score. This approach has been successful in facilitating the assessment of similarities and differences between EF tasks (Jurado & Rosselli, 2007; Miyake et al., 2000), however, the adoption of a task-focused approach for defining individual EFs assumes synergy between the outcome score of the task and the Executive Function in question (Jurado & Rosselli, 2007). This assumption is abundant within the literature as emphasised by the 98 different measures that have been used to establish the 68 different terms for EFs (Packwood et al., 2011), and has resulted in the overwhelming inconsistency of information, and likely contributes as one of the more significant causes of heterogeneity found in EF literature (Jurado & Rosselli, 2007).

A modern approach (e.g. Anderson et al., 2011; Miyake et al., 2000, 2012) that has served to inform many of the latest conceptual and theoretical models of EF today is to perform factor analyses *between* neuropsychological tests. A prominent body of research now exists that has demonstrated success using this method. This process operates by partialling out unique and shared variance between different outcome scores on EF tests. Unique variance between a set of outcomes scores is then used to validate the existence of a separable function that a set of given tasks are considered to measure.

Often, the unique variance between the set of outcome scores is only low to moderate, with non-unique shared variance demonstrating statistical significance in many circumstances (Miyake et al., 2000). However, this shared variance lacks operationalization and utility when not directly measurable or accounted for by an outcome score. Thus, it is often attributed to statistical noise, error, or non-executive processes (Chen et al., 2014; Miyake et al., 2000), or as a unitary construct of EF (Common EF) (Friedman & Miyake, 2016). The unique variance relating to the set of outcome scores then takes prime position and is used to demonstrate the existence of discrete EFs. This outcome can be attractive to the researcher as it provides a set of separable EFs that can be measured. However, the validity of this method is necessarily attached to the measures that were used, and the outcome scores that were chosen.

A recent review of latent variable derived modelling of EF by Karr et al. (2018), focused on the Unity and Diversity Model since the three EFs Inhibition, Updating and Shifting of this model were most commonly evaluated amongst the literature. Karr et al. (2018) found that the focus of these studies largely failed to move beyond the evaluation of the three factors. They further found that many researchers opted to design their projects around the three-factor model, instead of measuring EF as a larger construct and finding a three-factor solution from statistical modelling as an organic outcome (Karr et al., 2018). Although somewhat speculatory, Karr et al. (2018) suggested that this may be attributed to convenience and the popular attraction of the model to peer-reviewers. Consequently, a substantial body of research exists that appears to support the three-factor model of EF, albeit indirectly. Karr et al. (2018) also found that studies that aimed to replicate the three-factor model mostly carried out conceptual replication in place of direct replication in that alternative test batteries and varying outcome scores were used between them. Moreover, many of these used a low number of indicator variables and achieved poor statistical power, thereby straining the conventional assumptions of statistical modelling.

While latent variable analyses are providing new ground to increase our understanding of EF, it is imperative to be mindful that statistical factors are labelled by the researcher, and that they simply reflect the notion that a given outcome score shares a significant portion of variance with a separate outcome score. It is appealing to consider factors as mutually exclusive of one another, and certainly clinical assessments often subscribe to such EF modelling. However, the implication of distinct boundaries is artificial given the degree of researcher control in factor analysis (FA) procedures. The outcome of any Factor Analysis of EFs will depend on a) the number and diversity of EF tasks administered, b) the method of factor extraction, and/or c) the type of factor rotation. This therefore creates a circular argument where the question reflects the answer, and the answer implies the question. For example, a factor that is logically labelled "inhibition" is likely to be the outcome if the majority of tests included are known tests of inhibition. Thus, the mathematical derivation of a set of factors can only be as valuable as the diversity of tests included within it. The sample size required to successfully include a truly diverse range of clinical EF measures in Factor Analyses is ambitious, and likely a further contributor to the lack of consensus within the literature.

The current understanding of EF must focus on overcoming key challenges of EF assessment if data-driven accounts are to continue to inform, and subsequently validate theoretical constructs of cognition. To achieve this, two longstanding issues with EF tasks must be overcome, namely the ability of test performance to represent the real-world environment (Ecological Validity), and their ability to measure specific EFs in isolation (Task Impurity) with strong psychometric validity.

The issue of poor ecological validity has long-plagued EF assessment, as the translation of performance in lab-based assessment tasks to real-world performance yields very poor psychometric evidence (Burgess et al., 1998; Burgess et al., 2006; Royall et al., 2007). Given the nature of EF and its broad application to skills considered critical to activities of daily living, any given EF task must be able to predict everyday behaviours (Spooner & Pachana, 2006). However, many instances have been recorded where an individual's performance on tests of EF is inconsistent with their everyday abilities (Chaytor et al., 2006). Moreover, patients with frontal lobe lesions can perform equal to controls on traditional neuropsychological tests, yet still experience difficulties in everyday life (Chan et al., 2008).

It is therefore contended that assessment of performance in any context must consider the demands of the environment inherent to a particular task, at least to the same extent as to the skill that is being assessed. For example, inhibition if considered as a stand-alone Executive Function, cannot be considered to activate the same specific neurological regions, and to the same extent, when being asked to inhibit speaking in a library in comparison with inhibiting driving through a red light.

Despite clear indications that EFs cannot be wholly captured by a single outcome score, vary greatly between tests purported to measure similar constructs, and demonstrate consistently poor ecological validity, researchers continue to adhere to data-driven assessment outcomes. Furthermore, task impurity remains a significant problem for EF assessment

(Huizinga et al., 2006; Miyake & Friedman, 2012; Snyder et al., 2015). Failing to account for the multidimensionality or 'impurity' of EF measures risks a potential underestimation of important cognitive abilities that contribute towards performance compared to what is reflected in aggregated composite scores. Current measures either lack administration procedures to establish varied environmental demands which manifest separable EFs, or have scoring procedures that do not align with an administration that adequately distinguishes these environmental demands. For example, the interference score calculated from completion of the Stroop Test is mostly inferred as a measure of inhibition (Cothran & Larsen, 2008; Ropovik, 2014; Westerhausen et al., 2011), and thus a poor score is used to infer that an individual may have poor inhibitory abilities. However, the same interference score is also purported to provide an evaluation of working memory (Kane & Engle, 2003) cognitive flexibility (Zalonis et al., 2009), set-shifting of attention (Testa et al., 2012), response selection (Bender et al., 2016) and impulse control (Peterson et al., 1999). A deficit in any of these aspects of cognition could be reflected in the outcome score of the task, which may be branded as an inhibitory deficit, but little understanding afforded as to why this deficit is presenting in this testing environment.

Chapter 3

Understanding Task Demands

The notion of task impurity arises because efficient performance on any singular higher-order test of EF is underpinned by the integration of multiple skills. That said, researchers have relentlessly pursued the goal of isolating EFs for the purpose of assessment, and have largely denied the greatest influence on performance, namely the complexity and the novelty *within the task*. An argument may follow that such a pursuit would deny the nature of EF as a multi-componential higher-order skill *set*. However, when EFs are interpreted singularly, there is little clarity in theoretical advancement or clinical diagnosis. The EF tests that suffer less from criticisms of task impurity, and therefore demonstrate stronger psychometric reliability and validity, are those that isolate sub-components of task performance. They often begin with a sub-task that measures less demanding features of performance to scaffold and aid the interpretation of the higher-order demand.

For example, the Stroop Test provides conditions within its administration that enable direct comparison between trials (e.g. reading colour naming trial vs. incongruent colour/word trial). Although many versions of the Stroop Test stimuli and scoring exist, the interference score that is commonly derived ultimately reflects performance differences between a simpler and more familiar environmental demand (reading colour naming trial), and a more complex and novel environmental demand (incongruent colour/word trial). During the complex incongruent trial, (a) the tasks physical stimuli changes, (b) the instructions change from the previous trials, (c) the contextual information changes whereby the suppression of the automatic response of reading is required. In alignment with cognitive control research, these changes in task demands would require the recruitment of increased cognitive control for successful completion. By adopting an administration and scoring approach that both separates and then compares performance to other less complex trial demands, performance during the more complex trial of the Stroop Test can be observed. Results provide empirical insight into an individual's capabilities during less complex task demands (e.g. colour processing, reading comprehension, visual semantics (Strauss et al., 2006), prior to altering the complexity of the task to elicit the EFs necessary for successful completion. This outcome likely underscores the popularity and frequent uptake of the Stroop Test as a measure of EF within the literature.

A similar administration and scoring structure also exist within the popular Trail Making Test (TMT). The contrasting trials of the TMT enable a clearer delineation of changing task demands and an outcome score that corresponds to the impact of this change on performance. The TMT consists of two trials. The first trial (TMT-A), presents a series of numbers distributed across a page. The individual must draw a continuous line between the numbers in sequential order. Traditionally, the TMT-A is considered a measure of focused attention, visuospatial sequencing and scanning abilities (Salthouse, 2011; Sohlberg & Mateer, 2001; Strauss et al., 2006). The second trial (TMT- B), presents a task environment change whereby the individual must alternate between connecting numbers and letters in numerical or alphabetical order (e.g. 1, A, 2, B, 3, C, etc.). This change in task environment parameters, instructions, and available action rules all contribute towards the additional complexity that that is introduced during the TMT-B. The ability to contrast this performance with the earlier, less complex TMT-A, may explain its popularity as an assessment of the EFs of processing speed, set-shifting, cognitive flexibility and divided attention (Kortte et al., 2002; Lezak et al., 2012; Salthouse, 2011). The capacity to contrast complexity within task conditions as long demonstrated by the Stroop Test and TMT hints at the potential importance of being able to identify specific task environment changes that may influence performance during the more complex conditions.

3.1 Adapting Traditional Administration & Scoring Procedures

3.1.1 Fluency Paradigms

Measures of Verbal Fluency, as exemplified by the Controlled Oral Word Association Test, also known as the FAS Test (FAS), are considered to involve EFs such as attentional control, cognitive flexibility, information processing, inhibition, working memory, and planning (Abwender et al., 2001; Lezak et al., 2012; Strauss et al., 2006). The traditional administration of the FAS Test consists of three trials, each assigned to the letters 'F', 'A', & 'S'. Participants are asked to name as many words beginning with the letter for 60-seconds while following set rules (no proper nouns or repetition of multiple words using the same stem with a different suffix). Traditional scoring for the FAS Test includes the total number of acceptable words produced for all three letters (Ross et al., 2007; Shao et al., 2014; Tombaugh, 1999), or mean scores for all three trials (Barry et al., 2008).

The inherent assumption with traditional scoring procedures of the FAS is that all trials should bear conceptual similarity in their environmental demands. Thus, the EF requirements remain unchanged throughout the duration of the task. However, when fragmenting the traditional total scores into 15-second quartiles for each letter, Venegas & Mansur (2011) found that performance on the first 15-second quartile was significantly higher in comparison to other the other 15-second quartiles within the 60-second time limit for each letter. It was suggested that these performance differences were due to the support of semantic memory during the initial generation of words for each letter (Venegas & Mansur, 2011). Subsequent to the first quartile, complexity increases when semantic memory resources become exhausted, and the EFs of planning and monitoring of performance are required in order to generate further correct words that are free of repetitions or intrusion (Venegas & Mansur, 2011). This scoring approach demonstrates useful insight beyond traditional composite scores of the FAS, whereby the first 15-seconds for each letter represents familiar and simple demands to access the most well used or learned concepts from semantic memory stores, with more complex demands required to generate ideas while observing rules during the remaining 45-seconds. Complexity within this task may be attributed to a decrease in the number of alternative response choices available from a person's personal lexicon with each new word produced, the need to generate search strategies using known schemata (e.g. alliteration, category), and the ongoing retention and management of episodic information to both maintain the rule set, and to monitor words that have previously been provided.

3.1.2 Tower Paradigms

Tower Tasks are often used to establish complex, goal-directed environmental demands. Common Tower Tasks such as the Tower of London (TOL) or Tower of Hanoi (TOH) present a paradigm where a series of balls or discs are moved from a set starting position to a goal position, while following a set of instructed rules. The start and goal configurations change with each trial, creating a change in both the action rules available and the contextual information, with the common set of instructed rules remaining stable (e.g. a small disc cannot be placed on top of a larger disc). Due to the overall goal-directed and strategic nature of Tower Tasks, they are frequently used as a useful assessment of planning ability due to the inherent requirement to look ahead prior to enacting a sequence of 'planned' actions towards a specified goal state (Bishop et al., 2001). However, due to the overall complexity of Tower Tasks, they are also purported to require working memory, cognitive flexibility, strategic organisation, efficiency and goal setting abilities for completion (Bishop et al., 2001; Bull et al., 2004; Zook et al., 2004).

The Tower Task paradigm consists of multiple trials that are administratively different by the number of discs/balls to be moved, and/or by the number of moves required to reach the goal configuration. While each trial is administered independently, traditional outcome scores of tower tasks are often singular composite scores. These scores reflect either the total number of moves across all administered trials (Asato et al., 2006; Miyake et al., 2000), the number of successfully completed trials (Bishop et al., 2001), the number of moves used to reach a successful solution beyond the minimum necessary for that problem (Berg et al., 2010; Zook et al., 2004), or a point allocation for each correct trial summed to create a total score (Humes et al., 1997; Oosterlaan et al., 2005; Salnaitis et al., 2011). When compared to normative data (where available), this task may also provide an overall description of how *efficient* performance is based on the average score calculated across all administered trials. However, beyond this indication of efficiency, these traditional summated outcome scores for the Tower Tasks do not capture the various changing complexities and novelty demands within the trials that the paradigm demands, which have been demonstrated to hold significant influence over performance.

A traditional approach for attributing the complexity of task demands within Tower Tasks is to compare trials that vary in the number of minimum moves that are required between the TOH *start* and *goal* configurations. Thus, trials that require an increased number of moves are assumed to be amongst the more complex trials. While this approach holds value in the likely provision of increased availability of action rules, and an increase in the overall abstract relationships that need be considered for completion, attributing complexity solely to the number of minimum moves required is not considered to be a reliable rule-of-thumb (Kaller et al., 2011).

To further understand the demands within Tower Tasks, researchers have focused on the problem structure that exists between the start and goal configurations (Donnarumma et al., 2016). This approach considers both the steps required and the computational planning requirements needed to reach the goal state. For example, during a Tower Task trial the individual is required to engage in temporal abstraction to formulate sub-goals, which usually consists of freeing the largest disc/ball so it can be placed in its goal position first (Donnarumma et al., 2016). Once this subgoal is achieved, the individual is then free to formulate the next subgoal, which varies in quantity depending on the trial parameters and the location of each disc/ball (Kaller et al., 2011). An increase in the number of overall moves required provides a larger problem space where subgoals can exist. However, this relationship is not always linear, as a lower number of physical moves can sometimes encompass more sub-goals than a trial with a higher number of moves (Kaller et al., 2011).

The physical start and goal configuration of the apparatus for each trial has also been found to influence overall tower trial complexity. For example, Welsh & Huizinga (2005) demonstrated that during trials with a larger number of moves, goal configurations that require a stacked tower of disks (tower-ending) tend to be solved more accurately in comparison to configurations where disks are spread across two or more pegs (flat-ending problems). Reasons for this difference have been attributed to the occurrence of a potential goal recursion strategy, where tower-ending tasks share similarity in their final configuration (Welsh & Huizinga, 2005). The use of a goal recursion strategy, whether explicit or implicit, can guide the person to achieve sub-goals in the correct order, and thus improve their performance accuracy (Welsh & Huizinga, 2005). Therefore, two trials of the same number of moves or discs may be establishing different environmental demands; one that is somewhat familiar and reliant upon schemata and episodic information gained from previous trial exposure, and one that is more demanding of cognitive control due to the novelty of configuration.

Kaller et al., 2011 demonstrated that additional parameters that can influence TOH complexity and performance can also include; *search depth* (the number of intermediate moves that have to be considered before the first sub-goal can be reached); the number of *counterintuitive moves* (an intermediate move that requires the movement of a disc that is already on its target peg); and number of *suboptimal paths* (a goal solution is reached in more than the minimum number of moves). Although, it remains unclear which of these parameters are the most influential contributor to Tower Task complexity, this insight further exemplifies the need to consider the complexity of task demands within the trial parameters of Tower Task when used as a measure of EF. Due to the novel, complex, and also potentially familiar conditions that can vary between Tower Trials, an understanding of strengths and weaknesses during these varying demands may be underrepresented using traditional administration and scoring procedures.

3.1.3 Maze Paradigms

Hidden Maze Paradigms have also been used in various forms to establish complex environmental demands for testing EFs and cognitive control during goal-directed behaviours (Lezak et al., 2012). The original Austin Maze (AM) test, and now the recently licensed digital version the Groton Maze, present a 10x10 grid of tiles whereby individuals are required to tap each tile to learn a hidden pathway between a start point and an end point while following rules. The rules can include only being able to move one tile at a time, and only being able to move up-down, left-right and not diagonal. Feedback is provided by a sound and/or a corresponding green/red stimulus to confirm whether the hidden path is being correctly identified. This feedback is applied across up to 10 attempts until error-free trials are achieved.

Overall, Hidden Maze paradigms are considered to require spatial working memory and retention of episodic information regarding rules, which ultimately rely on the cognitive processes of attention, visuomotor processing speed and integration, and decision-making (Darby & Walsh, 2005; Pietrzak et al., 2008). Due to the task environment demands that are established, researchers also utilise Hidden Maze paradigms to a measure feedback utilisation, planning, goal setting, and inhibition abilities (Bowden & Smith, 1994; Tucker et al., 1987).

Scoring alternatives for Maze Tasks include the number of trials taken to reach two or three sequential error-free trials (Crowe et al., 1999), total number of moves and/or total time to complete all trials (Pietrzak et al., 2008), or the total number of errors across all trials (Crowe et al., 1999). However, the demands across each trial may not necessarily be comparable. It has been proposed that early trials of the AM test may reflect predominantly visual-spatial abilities while the individual orientates themselves to the path (Crowe et al., 1999), whereas later trials reflect the visual-spatial recall of the path (Crowe et al., 1999). However, the composite tallying or averaging of performance across all trials may potentially negate the very nature of the changing demands of the task and its administration. The changing demands within the Hidden Maze tasks provide the central ethos to its very purpose for understanding how people discover, adapt and later recall a hidden path following a set of rules. Yet, such detail and understanding of how individuals perform within these changing demands currently remain limited by composite scores that inherently assume parallel demand between trials.

3.1.4 Block Design Paradigms

Successful performance during the Block Design (BD) task has been documented to require multiple cognitive and executive abilities. Arguably, the most widely utilised BD task is the version within the Wechsler Adult Intelligence Scale (WAIS) (Wechsler, 2008), which is a derivative of the original Kohs (1923) Test. The BD test has been associated with EF and frontal lobe functions and is considered to assess strategy and planning skills in the context of visual-spatial construction (Brown et al., 2012; Lezak et al., 2012; Wallesch et al., 2001). For this test blocks with various full or split-half coloured faces (red, white, or red and white) are required to be manipulated to replicate geometric designs as displayed in a stimulus book. As the test progresses, the designs increase in complexity, either in the number of blocks required to be used or the orientation of the overall design, and provision of supporting structures.

Performance is largely found to be negatively impacted when patterns have diagonal blocks present, and overall perceptual and task uncertainty is increased. Task uncertainty and perceptual cohesiveness are found to affect the level of local or global figure processing and the overall complexity of the tasks (Cardillo et al., 2017; Rozencwajg & Corroyer, 2001). For example, if the design requires the selection of a full-sided block, the participant is only required to decide whether it is red or a white face that requires selection (Royer et al., 1984). Alternatively, if the task required the selection of the split-half red/white block face, task uncertainty is increased due to the four different configurations the block could possess. Task uncertainty was also found to manifest due to the overall pattern of design that is required to be constructed. Performance is found to be similar between four and nine block designs if all are limited to the use of solid-face blocks. However, marked performance differences can be found when design patterns require the use of split coloured blocks (Royer et al., 1984). That said, there is a component of ‘nesting’ embedded across all trials, as smaller patterns of block construction can be integrated within other patterns to form the overall target stimulus.

The nature of these demand effects is somewhat accounted for within the WAIS BD administration (Rozencwajg & Corroyer, 2001). The increase in complexity is mapped to nine block trials that are afforded a longer time for completion in comparison to four block designs. Whilst this provides for the control of psychomotor response in overall total outcomes scores, this method of scoring does not provide clear insight into performance during different demand representations. Moreover, it remains unclear whether performance during each of the different levels of complexity within the block design can significantly influence performance, or if performance across the task is equitable between trials.

3.2 Neurological Representations of Task Demand

Behavioural performance differences that can be quantifiably measured *within* trials of an assessment task are only available when their scoring and administration procedures are adapted accordingly. While it may be apparent where these demands may exist, for example, a change to instructions, actions rules, or the retention and utility of information, some tasks do not provide an explicit demarcation. Many neuroimaging studies focus their activation contrasts *between* a set of neuropsychological tests, often due to the research aiming to draw comparisons *between* their purported cognitive skillsets. However, neuroimaging studies that have captured activation amplitude *within* various trials and stages of EF tasks offer a supplementary source of understanding within task demands.

3.2.1 Complexity Dependent Activation during Working Memory Paradigms

WM tasks require the retention of information 'in mind' whilst higher-order problem solving can be undertaken. In the example of a sentence reading task, cognitive control network activation patterns are found to be complexity dependent between sentence reading and recall, and by combining the two (read-recall) as a dual-task (Bunge, 2000). Across these tasks, the LPFC, right PFC, left middle temporal gyrus, and bilateral anterior cingulate regions were all activated (Bunge, 2000). During recall and dual-task conditions, additional activation within the bilateral parietal, occipital and cerebellar regions was observed, with greater activation found under the dual-task paradigm. Interestingly, during the most complex condition (dual-task) no novel activation outside of those previously activated by the reading and recall tasks alone were reported (Bunge, 2000). Instead, an increase in activation amplitude within these regions was observed as the gradient of task demands increased from that of sentence reading, to remembering, and finally to the most complex condition, reading and remembering (dual-task). The absence of additional recruitment of activation across the cortex during dual-task demands demonstrates that the cognitive resources allocated are provided by the networks subsumed by the two tasks separately. The involvement of the PFC across all task conditions highlights the need for cognitive control, and the upregulation of the neural regions identified demonstrates that cognitive demand at various stages of a task can be structurally analogous, and activity dependent.

Similar cognitive demand activations are found when increasing the number of consonants for encoding and retrieval from WM. During a Sternberg WM task, Michels et al. (2010) demonstrated a task load dependence of neural activation within the frontoparietal network -predominantly within the anterior cingulate cortex (BA 32), posterior cingulate cortex (BA 23/31), medial pre-frontal cortex, and posterior parietal cortex. Fronto-parietal networks have also demonstrated a similar load-dependent activation when sub-goals are added to WM tasks (Farooqui et al., 2012).

During an alternative measure of WM, the n-back paradigm, task load dependency was found to account for the percentage of increase in amplitude within the DLPFC, superior parietal lobule, and precentral gyrus (Jaeggi et al., 2003). The n-back conditions included a less complex 1-back condition that was compared to a more complex 3-back condition (Jaeggi et al., 2003). The increased activation reported within the DLPFC across WM demands is thought to be in response to increased maintenance and retention demands of the task (Manoach et al., 1997). In turn, this can be further attributed to the requirement of the PFC to exert cognitive control as WM demands increase.

3.2.2 Complexity Dependent Activation during Planning Paradigms

Activation of the DLPFC is considered to be complexity dependent during completion of the TOL (van den Heuvel et al., 2003). Conversely, the posterior parietal and occipital cortices are considered to be complexity independent (Dagher et al., 1999). An early study by Baker et al. (1996) demonstrated that different stages of the TOL task lead to different neural responses. Analyses of performance scores revealed that the length of response times and the number of incorrect responses increased as the task conditions changed sequentially from two to five moves (Baker et al., 1996). During earlier trials (2-3 moves), the authors reported extensive activation in the insula/operculum, with activation primarily bound to the RLPFC (Baker et al., 1996). In contrast, during later conditions (4-5 moves) an increased magnitude of activation was recorded in prefrontal, premotor and medial parietal cortices. Greater frontal activations were restricted to the right DLPFC, Premotor Areas (BA 6), the left DLPFC, and right rostrolateral prefrontal cortex, and decreased activation found in the insular/opercular area

(Baker et al., 1996). Subsequent research also identified that cingulate, basal ganglia, thalamic and cerebellar regions become significantly more active during six-move TOL trials in comparison to two-move paradigms (Unterrainer et al., 2004).

Similar to the activation patterns recorded during various stages of WM tasks, the recruitment of rostral frontal regions and the upregulation of activity within posterior regions during a task predominantly of planning demonstrates that cognitive demands within the different stages of tower tasks can vary considerably. The necessity of this upregulation as task demands increase was emphasized when better performance outcomes on the TOL were found to be associated with increased activation of the right DLPFC, right superior temporal regions, and right inferior parietal regions (Unterrainer et al., 2004).

As previously mentioned, what makes some of the TOL conditions more complex can be attributed to the number of moves required to reach the goal state (Welsh & Huizinga, 2005), but is also dependent on the need to generate sub-goals, and the search depth required for completion in the minimum number of moves possible (Berg et al., 2010; Kaller et al., 2004; Spitz et al., 1982). Search depth refers to the number of intermediate moves one must consider prior to reaching the first goal move (Kaller et al., 2008). Neurologically, increasing search depth in the TOL also demonstrates a hierarchical rostro-caudal activation pattern, as increased activations in anterior prefrontal regions are found in response to the increased cognitive demand (Kaller et al., 2011).

3.2.3 Complexity Dependent Activation during Cognitive Flexibility Paradigms

The Wisconsin Card Sorting Test (WCST) is considered a measure the EFs; updating, set-shifting and cognitive flexibility (Lezak et al., 2012; Strauss et al., 2006), and is known to tap into the neural networks of WM (Konishi et al., 1999). Similar hierarchical cortical activations are also found during different stages of the WCST. In an effort to understand the cortical representation of the WCST, Lie et al. (2006) investigated its neural correlates by manipulating the demands of the task through varying the test parameters from a less complex explicitly instructed matching paradigm, to a more complex non-instructed paradigm (Lie et al., 2006). Increased complexity within the WCST paradigm revealed a parallel increase in the magnitude of activation within the right PFC and caudal ACC (Lie et al., 2006). Common activations to all three task conditions were found in the posterior region of the caudal ACC, and left PFC, with signal intensity increasing as the task requirements became more demanding (Lie et al., 2006). Lie et al. (2006) considered this increase of neural activity to be associated with the diverse cognitive skills (e.g. error detection, feedback utilization, set-shifting, working memory and increasing attention) that are required for successful WCST completion.

More recently, Yoshida et al. (2010) developed and implemented the Multi Feature Sorting Task (MFST) which is fundamentally similar to the WCST but provides a hierarchical structure to allow for behavioural decisions to be better isolated. The task produces a higher-order meta-rule that requires the participant to search for a rule that belongs to another meta-rule class, and a lower-order rule that requires a switch within the meta-rule class (Yoshida et al., 2010). The design of the MFST enabled the authors to investigate the hierarchical processing and neural correlates of higher-order vs. lower-order rule switching and updating. During the lower-order condition, the right PFC was activated, and during the higher-order condition, increased activation was recorded in the left DLPFC and left posterior PFC (Yoshida et al., 2010). The anterior cingulate cortex was the only cortical region activated across all conditions (Yoshida et al., 2010). The consistent activation of the anterior cingulate cortex was considered to be in response to error detection requirements across all tasks, whereas the DLPFC was considered to be involved in the decision making and coordination of information stored in episodic memory (Yoshida et al., 2010). Similarities exist between the activations reported during split conditions of the WCST (Lie et al., 2006) and MFST, particularly the increased recruitment of regions within the PFC. Arguably, the differences between the additional neural regions captured by each study are representative of the variations in the specific cognitive demands of each overarching task.

Chapter 4

Rationale and Aims

The neuroscientific research discussed thus far highlights a number of fundamental considerations. Firstly, the upregulation of posterior regions along with the PFC in response to increased cognitive demand demonstrates the diverse range of cortical networks that can be called into service during EF task completion. Secondly, the nature of this activation is likely hierarchical, with a rostro-caudal organization demonstrated by the upregulation of the PFC regions when task demands increase that require cognitive control. Importantly, this hierarchy can be observed across multiple classic neuropsychological measures of EF that tap into a variety of its behavioural constructs. Finally, the up-down regulation of this network in response to the changing conditions within and between different tasks demonstrates its adaptive nature to suit the demands that are required of it.

Collectively, the nature of these cortical activations during traditional EF tasks provides a preliminary insight into the task-dependent nature of EF and its relationship to cognitive control demands. Although not exhaustive, this collection of research demonstrates the neural interdependence and dependence that neural regions share to allow successful completion of complex, goal-directed tasks. Evidence suggests that pathologies impacting these activations can be detrimental to various aspects of EF performance across a variety of ages and clinical syndromes (Dickstein et al., 2006; Durston et al., 2003; Gilbert et al., 2008; Minzenberg et al., 2009; Sebastian et al., 2012; Vaidya et al., 2005). Thus, the nature of this overall task-dependent activation must be considered during the clinical assessment of EF.

Albeit via different modalities, neurological and behavioural investigations of EF serve to inform the same cognitive construct. The task demand approach adopted during neuroimaging studies has enabled a greater insight into the functional neural correlates that various EF tasks can elicit. As our theoretical representations of EF are often informed by the tools and outcome scores that we attribute to measure it, it is imperative that we continue to maximize our understanding of the performance information that is available from these tasks. The neuroscientific evidence provided here offers a progressive avenue towards how the future interpretation of existing EF tasks can be augmented to further maximize our assessment of an individual's response to task demands. For example, changes in performance outcome scores that are seen to parallel the increased activations patterns are able to provide a behavioural measure of task demands. Thus, the two methods of enquiry are representative of the same demands that are required.

Cognitive control research has provided an understanding of the key features of task environments that form task complexity to influence the overall upregulation of control requirements. This research may offer a meaningful avenue to further uncover what demands within EF tasks exist to better inform neuropsychological assessment outcomes. The current absence of administration procedures and outcome scores that reflect these demands highlight the limitation of current standardized scoring procedures to allow for, or respect changes in, the complexity that can occur within a task. Therefore, a representation of task demands from between *and within* various EF tasks must be included amongst theoretical representations of EF if our understanding of the construct, and its assessment, is to be elucidated.

This circular relationship between neuropsychological tests informing theoretical constructs and vice versa was discussed by Barkley (2012) who further suggested that current neuropsychological measures only allow for an understanding of the components of EF, but not its adaptive nature. However, the neural research reviewed here supports the advantage of breaking down different task conditions to observe an adaptive EF that is responsive to task complexity and novelty. This requires shifting the focus from only variance between tasks, to include within tasks. The measures themselves must be divided into conditions that reflect their level of complexity and novelty, making them more consistent with neuroscientific EF imaging paradigms, and cognitive control theory.

Interpreting outcome measures based on both within and between task performance, through the lens of cognitive control demands, will arguably facilitate convergence of the theoretical and neuroimaging evidence that is long overdue. The impetus to redirect research focus in this way is substantiated by various lines of evidence. The MD system and DMN theories support the proposal that the rostral upregulation of neural networks and additional regional activation specific to task requirements demands consideration of both between and within EF task performance. During less complex tasks the neural networks that are active likely represent the foundational requirements of the task, which as previously proposed, may reflect attentional control requirements (Anderson et al., 2011a) or a common EF factor (Miyake & Friedman, 2012). As task difficulty increases, additional task-specific activation occurs in parallel to foundational systems, demonstrating the importance of integrated networks in successful task performance.

Failing to consider the complexity and novelty within a task will only serve to perpetuate the long-standing issues of attempting categorization of EF domains, without adding a great deal of progress to overall theory. What remains to be understood is the *extent* to which complexity and novelty, and thus cognitive control and EF, are present at various levels of within a task. Friedman & Miyake (2016) acknowledge this in their work by placing their unity/diversity framework at an intermediate level of complexity, and by excluding higher-order planning tasks that typically obscure mathematical outcomes because of increased shared variance in complex task performance. Their work represents an exciting beginning, but the exclusion of an understanding of the complexity of demands whereby EF may be required across different tasks has restricted their theory to include only three domains, and more recently, the common EF factor. Without determining which task conditions can be classified in regards to their complexity and novelty, the understanding of EF involvement during various task demands falls short.

There is an implicit call by the research collective for convergence of related disciplines of cognitive neuroscience, EF theory, and neuropsychological practice. Much like the work of Atkinson & Shiffrin (1968) and their seminal modal model providing a storage perspective of memory, it was the understanding of levels of processing initially purported by Craik & Lockhart, (1972) that advanced memory theory to include both function *and* mechanism. The current EF literature is at a similar precipice, in which cognitive theories have demarcated the skills that are subsumed under the overall construct, and neuroimaging studies are working toward understanding the nature of their interactions. Cognitive control research has demonstrated that environmental demand is best captured when considering the internal requirements of a task that reflect the gradient of hierarchical controlled cortical activation required to solve a specific overarching problem set.

By seeking to capture the complex and novel demand features both within and between multiple measures of EF, an overall insight will emerge into the degree of demand that is required by an individual task, and the potential commonality of complexity and novelty between different measures of EF. This approach demands a re-conceptualisation of the administration and scoring of traditional cognitive measures to capture the levels of demand within tasks. In doing so, only then can insight be gained towards whether a particular trial within a task is under similar complex and/or novel demands and therefore best reflected by their traditional composite outcome score, or whether these demands vary within a task that are best represented by multiple outcome scores. The ultimate goal of this approach is to bring further clarity to EF assessment and theoretical accounts by capturing the variability of demands during the most complex tasks where skill overlap has impeded progress.

4.1 Aims

This thesis aims to reconceptualise traditional scoring approaches of a variety of neuropsychological tests through the lens of Cognitive Control Theory. Essentially, this thesis will disregard traditional EF nomenclature and assess each established test according to the demands required for its execution. Such an approach necessitates an operationalisation of key demand features that to date is not provided by the cognitive control literature. This approach therefore requires the development of a Demand Classification System (DCS) that functions to quantitatively appraise the complexity and novelty demands inherent within a task. Specifically, the appraisal will consider the degree of its *complex* demands, (1) the level of abstraction required, (2) the contextual information of task parameters, (3) the number of possible response combinations available, (4) the number of instructions and rules, and (5) the dual nature of the task. In addition, the appraisal will address tasks' *novel* demands, (6) automatic and controlled behaviour required, and the required integration and processing of (7) schematic, and (8) episodic information that is internally task-centred and/or from previous knowledge and experience.

In an effort to address the aforementioned, the overarching research question is whether demand on cognitive resources can be quantified using performance on neuropsychological tests? This question was investigated by integrating Cognitive Control research and EF literature to:

(a) Operationalise a Demand Classification System (DCS) in an attempt to quantify performance according to task complexity or novelty.
(b) Apply and mathematically validate the DCS to tasks identified as producing single outcome scores, even though research has arguably demonstrated that successful performance is characterised by more than one skill.

(c) It is also necessary to ensure that test elements classified as similar in complexity and novelty hold together mathematically. This will allow comparisons not between tests but across demands, which will be termed a Global Demand Classification (GDC).

(d) Develop a model to depict the relationships between each of these GDCs.

Chapter 5

General Method

5.1 Sample

A total of 105 participants aged between 18 to 55 years (*M*= 30.00, *SD*= 7.11) were recruited via convenience sampling after responding to a flyer that advertised the project in the community and university setting (Appendix A). The current project aimed to generalise any findings to the neurotypical adult population and therefore participants were eligible for the study if they met the following criteria:

a. Aged between 18-55 years at the time of participation.
b. Had no currently diagnosed or known neurological, intellectual or psychological illness.
c. Had not completed a neuropsychological test battery within the previous 24 months of the testing date.

The age range of the current study was selected in line with the body of evidence that supports the maturation and absence of age-related EF decline within this developmental period (Kray et al., 2005; Vasquez et al., 2016; Wiebe & Karbach, 2018).

To further ensure that the sample represented a healthy adult population, all participants completed of the Wechsler Abbreviated Scale of Intelligence – Second Edition [WASI] (Wechsler, 2010), which served as a screening tool for the presence of intellectual impairment. BD and Vocabulary subtests of the WASI were completed to produce a calculated Full-Scale IQ- 2 (FSIQ-2) score (Sattler & Ryan, 2009).. FSIQ-2 scores within the range of 71- 130 were considered to represent the absence of intellectual impairment or exceptionality (Wechsler, 2010). Participants with FSIQ-2 scores outside of this range were excluded from further analysis. In total, two participants were omitted from the study due to their FSIQ-2 scores falling outside of the inclusion range. The final sample size was *n* = 103, with the *M* and *SD* of age remaining unchanged.

As displayed in Table 2, the mean and standard deviation of FSIQ-2 scores for the current sample were slightly above (for *M*), and slightly below (*SD*) population norms (*M*= 100, *SD*= 15) as reported in the WASI manual (Wechsler, 2010). However, as the FSIQ-2 average fell within 1 standard deviation of the population norm, the sample was considered an acceptable representation of the intellectual functioning of the general population.

Table 2

Descriptive Statistics for Sample Performance on the WASI

	M	*SD*	Range	Skewness	Kurtosis.
FSIQ-2	108.97	10.29	84-130	-0.170	-0.511
Vocabulary T-Score	53.87	8.16	33-71	-0.197	-0.527
Block Design T-score	56.67	7.27	38-71	-0.192	-0.429

Note. FSIQ-2 = Full Scale IQ- 2

A large portion (43.7%) of the sample were aged between 26-30years (Table 3), with 70.9% of the total sample being females (*n*=73).

Table 3

Age Range and N Distribution of the Current Sample

Age Group	*N*	%
18-25 years	24	23.3
26- 30 years	45	43.7
31-35 years	13	12.6
36-40 years	10	9.7
41-45 years	6	5.8
50-55 years	5	4.9

Note. Total sample size *n*=103.

5.2 Measures

A large battery of neuropsychological measures was administered to capture performance across a wide range of EFs and cognitive skills in order to address the current aims.

Test selection for the current study was based on the following criteria:

a. A widely accepted, reliable and valid measure of EF within the literature
b. Multi-faceted in demand requirements and allowing for the scoring of individual trials, OR;
c. Singular in demand requirements, and with a total score that is considered a valid representation of performance within the literature.

5.2.1 Wechsler Abbreviated Scale of Intelligence – Second Edition (Wechsler, 2010)

The WASI is used for screening and estimating an individual's current intellectual functioning (Wechsler, 2010). As previously stated, two WASI subtests were administered (Vocabulary and BD) which enabled the calculation of a FSIQ-2 score for each participant. The FSIQ-2 score is considered a highly reliable score, with reliability coefficients ranging from .90 to .98 (Groth-Marnat, 2003; Strauss et al., 2006; Tuokko & Hadjistavropoulos, 2014; Wechsler, 2010). Administration and scoring were conducted following the standardised instructions. Specifically, the Vocabulary subtest is a measure of verbal knowledge, expressive vocabulary, crystallised intelligence and general intelligence (Wechsler, 2010). The examiner presents four picture items and 38 word items in sequential order to the participant. The participant verbally responds to each word by explaining what a particular word is (*e.g. What is a Calendar?)* or what the word means (*e.g. what does famous mean?)*. The WASI BD subtest is a measure of general intelligence in the realm of spatial visualisation, visual-motor coordination, and abstract conceptualisation abilities (Wechsler, 2010). During administration the examiner presents to the participant a series of 13 geometric patterns from a stimulus book that progress in difficulty from a two-block configuration to a series of four, and then nine block configurations. Each block has two white sides, two red sides, and two split-half red and white sides. The participant is given the appropriate number of blocks for each configuration and is asked to replicate the pattern within a specified time limit.

Calculated WASI T-scores were summed from BD and Vocabulary subtests and converted to normed FSIQ-2 score based on a person's age provided with the WASI manual (Wechsler, 2010).

5.2.2 The Stroop Colour-Word Test – Victoria Version (The Stroop Test) (Spreen et al., 1998)

The Stroop Test is a frequently utilised EF measure of cognitive control and inhibition abilities. The task requires the individual to maintain a goal in mind, whilst suppressing a habitual response (Spreen et al., 1998; Strauss et al., 2006). The Victoria version of the Stroop Test was selected due to its brief administration time, and its ability to minimise practice effects (Lezak et al., 2012).

The Stroop Test is comprised of three trials including a *Colour* trial, a *Word* trial, and a *Colour-Word* trial. Each trial displays 24 items on a 21.5 x 14cm card, consisting of six rows containing four items each. The *Colour* trial presents a series of coloured dots (green/ blue/ yellow/ red) printed in a variable sequence. The participant is instructed to read aloud across the page the name of the colour of each dot as quickly and as accurately as possible. The *Word* trial presents a series of commonly used words (e.g. 'when', 'and', 'over', 'hand') printed in green, blue, yellow or red ink. The participant is instructed to read the word aloud as quickly and as accurately as possible. The Colour-Word trial presents the written names of colours (green/ blue/ yellow/ red) that are printed in a mismatched colour ink (e.g. the word 'blue' is printed is red ink). The participant is instructed to name the colour of the ink as quickly and as accurately as possible.

Test-retest reliability coefficients for the Victoria version of the Stroop Test are high, .90 (colour trial), .83 (word trial), and .91 (colour-word trial) (Strauss et al., 2006). Moderate correlations have been found within the test itself within a healthy sample (Pineda & Merchan, 2003), suggesting that each trial is requiring a similar, but not identical ability. For each trial, the time to complete was recorded along with the number of errors. Participants were instructed prior to the commencement of the test to self-correct any errors during each trial. For any error that was not self-identified, the participant was awarded a time penalty of 1-second. Total time to complete each trial was used as the outcome score.

5.2.3 Digit Span -Backwards (Darby, 2014a)

The Digit Span Backwards task provides a measure of verbal WM abilities (Lezak et al., 2012; Strauss et al., 2006). An Apple iPad version of the Digit Span task was implemented using ‘Span Tests’, Version 1.2 (Darby, 2014a). The examiner begins by reading out a series of four numbers (e.g. ‘5, 4, 1, 8’). Each number is read aloud at a pace of one number per second. After the investigator completes reading the sequence, the participant is then required to repeat back the numbers in the reverse order. If a correct sequence is recorded, the researcher reads aloud a new set of numbers that increases in span by one digit for every two consecutively correct trials. The test is completed when the participant fails to repeat all numbers correctly across two trials. Each correct trial was awarded 1-point, and summed to produce a Total Trials Correct Score the reliability coefficients for the Digit span tests range from .80 to .89 (Strauss et al., 2006).

5.2.4 Visual Span- Backwards (Darby, 2014a)

The Visual Span Backwards task provides an assessment of visuo-spatial WM ability (Smyth & Scholey, 1994; Strauss et al., 2006). An Apple iPad version of the Visual Span tasks was implemented using ‘Span Tests’, Version 1.2 (Darby, 2014a). The task is an electronic version of the commonly used Corsi Block Tapping Test (Kessels et al., 2000). The tasks begin by placing the iPad in front of the participant that displays a series of nine boxes on the screen. A digitized hand points to a series of boxes in a specified order. The participant is then required to tap the boxes in a reverse order. The sequence of boxes begins at three and increases by one additional box for every correct trial. The task concludes when the participant completes two trials incorrectly. Each correct trial was awarded 1-point and summed to produce a Total Trials Correct score.

5.2.5 Trail Making Test (TMT) (Reitan, 1955)

The TMT consists of two trials. The first trial, the TMT-A, is a considered as measure of focused attention, visuospatial sequencing and scanning abilities (Salthouse, 2011; Sohlberg & Mateer, 2001; Strauss et al., 2006). The second trial, the TMT- B, is a frequently used tool for the assessment of processing speed, set shifting abilities, cognitive flexibility and divided attention (Kortte et al., 2002; Lezak et al., 1995; Salthouse, 2011; Tombaugh, 1999).

The TMT-A required each participant to draw a continuous line using a pencil between 25 sequentially encircled numbers that were distributed across an A4 sized page. The TMT-B required participants to alternate between connect numbers in sequential order and letters in alphabetical order (e.g. 1, A, 2, B, 3, C, etc.). Any errors that were made across both trials were alerted to the participant, and they were asked to correct their response. The scores for each of the TMT tasks were calculated using the total time for each trial. The TMT demonstrates good test retest reliability for TMT-A (r=.79), and excellent test-rest reliability for TMT-B (.89) (Strauss et al., 2006).

5.2.6 5-point Test (Strauss et al., 2006)

The 5-point test is a popular measure of design fluency that requires EFs of working memory, short-term and delayed recall, inhibition, problem solving, and cognitive flexibility (Goebel et al., 2009; Lezak et al., 2012; Tucha et al., 2012). Participants were instructed to draw as many unique figures as possible within three minutes on a stimulus sheet that was provided. The sheet was A4 in size and consisted of 40 dot matrices identical to that of a 5-dot arrangement on a dice. Participants were instructed to follow a set of rules during the tasks, which stipulated that (a) only straight lines are allowed to be drawn, (b) all lines must connect dots on the stimulus page, (c) no figure can be repeated, (d) only single lines can be drawn, and (e) all lines must be continuous. Only correct designs that did not violate any of these rules were considered for scoring. Each correct design was awarded 1-point, and total scores were calculated for each 60-second increment. A good test-retest reliability for unique figures of the 5-point has been previously demonstrated (r=.77) (Tucha et al., 2012) and (r=.78) (Fernandez et al., 2009).

5.2.7 The d2 Test of Attention (Test of d2) (Brickenkamp & Zillmer, 1998)

The test of d2 is a pen-and-paper letter cancelation task that is a measure of attention, visual scanning, and speed of processing, and cancellation abilities. The task was selected for inclusion within the study due to the inhibitory requirements needed to rapidly attend to target items and filter out irrelevant configurations (Neill et al., 1995). The test comprises of 14-trials, with each trial containing a row of 47 “p” and “d” characters. Each character is configured with either one to four dashes that are either placed individually or in pairs above and/or below each letter.

Participants were instructed to cancel out target configurations as quickly and as accurately as possible. The target configurations included a letter “d” with either two dashes above, two dashes below, or one dash above and one below. Cancelation of the target configurations were to be done by moving left to right across the page, with a 20-second time limit per trial (line). Scoring included counting the number of correct cancelations across the 14 trials to provide a total speed corrected score. The Test of d2 has demonstrated excellent test-retest reliability (r= .90) previously (Steinborn et al., 2018).

5.2.8 Map Search (from the Test of Everyday Attention) (Nimmo-Smith et al., 1994)

The Map Search subtest is a cancellation style task that is traditionally used as a measure of selective attention (Nimmo-Smith et al., 1994). Participants are presented with a realistic colour map of an area of Philadelphia (USA). Participants are then presented with a stimulus card that displays a target symbol that denotes a mechanic. The participant was then asked to find the target symbols amongst other distractor symbols within the map. Testing time was 2-minutes in total, with the number of correct symbols scored for each 60-seconds of the task calculated (Nimmo-Smith et al., 1994). The maximum score achievable was 80 correct symbols across the 2-minutes. The number of correct symbols identified within 61-120-second period of the Map Search subtest was selected as the performance score for the current study. During this period the participant was assumed to have successfully identified a majority of symbols that were easiest to identify, and the demand of the task was therefore increased to require the use of additional cognitive skills to successfully continue to find the correct symbol amongst distractors. The Map Search subtest has previously demonstrated good Test-retest reliability (.80) (Nimmo-Smith et al., 1994).

5.2.9 Visual Elevator (from the Test of Everyday Attention) (Nimmo-Smith et al., 1994)

The Visual Elevator subtest is a measure of attentional switching and cognitive flexibility (Nimmo-Smith et al., 1994). Traditional administration of the Visual Elevator was conducted following the TEA administration manual. Participants were presented with a stimulus book that contained 10 trials. Each trial presented a visual depiction of a series of elevators that each represent elevator floors. Participants were asked to count in a sequential order the number of elevator images displayed, until a switch signal was encountered that was depicted as either an ‘up’ or ‘down arrow’. If an up’ arrow was presented, the participants were required to say out loud the word ‘up’ and continue to count from the number prior to the arrow presentation. If a ‘down’ arrow was encountered, the participants was required say the word 'down' and count in a reverse order from the number prior to the arrow presentation. The TEA timing score was used as the outcome score which was calculated by taking the total time for all correct trials, divided by the total number of switches in the correct items. The Visual Elevator subtest has previously demonstrated good Test-retest reliability (.79) (Nimmo-Smith et al., 1994).

5.2.10 Elevator Counting with Reversal (ECR) (from the Test of Everyday Attention) (Nimmo-Smith et al., 1994)

The ECR subtest was designed as an auditory equivalent to the visual elevator tasks. However, the ECR is considered to require additional cognitive resources due to the demands placed on verbal WM by the absence of a visual aid (Nimmo-Smith et al., 1994). During the ECR task, an audio recording is played that contains a series of 10-trials of elevator scenarios. The participant was required to track the elevation of building floors that the elevator travelled that were represented by a series of audio tones. This required counting of a string of mid-range tones (count tone) that represented one floor being travelled. A high-pitched tone (signal tone) was played during each string that served as an indicator that the elevator was moving upwards through the building, but the tone itself was not to be counted as a floor. This required the participant to count forwards after a high-pitched tone was played. A low-pitch tone (signal tone) was played during each string that served as an indicator that the elevator was moving downwards through the building, but the tone itself was not to be counted as a floor. This required the participant to count backwards after a low-pitched tone was played. The series of high-pitched and low-pitched tones varied in quantity across the 10 trials. The outcome score for the ECR was the number of correct items across the 10 trials, as per the standard scoring procedures within the TEA manual (Nimmo-Smith et al., 1994). The ECR subtest has previously demonstrated acceptable Test-retest reliability (.68) (Nimmo-Smith et al., 1994).

5.2.11 Tower of Hanoi (TOH) (Bishop et al., 2001)

The TOH task provides an assessment of an individual's overall planning ability and is found to require WM, cognitive flexibility, strategic organisation, efficiency and goal setting abilities (Bishop et al., 2001; Bull et al., 2004; Zook et al., 2004). The TOH presents the participant with a wooden apparatus that has three rods of equal height. The aim of the TOH is to transfer a number of flat, wooden disks that differ in size from a starting state to a goal solution. The TOH task is comprised of 13-trials including one two-disk trial, eight three-disk trials, and three four-disk trials. The examiner sets the disks into to their starting state and displays to the participant a visual image of the goal solution. The participant must move the disks using the minimal amount of moves possible to reach the goal solution. During completion, the participant is required to adhere to three rules, (1) only one disk may be moved at one time, (2) no larger disk can be placed on top of a smaller disk, (3) only the top disk on a stack can be moved first, prior to those beneath it.

During some traditional administrations of the TOH, particularly when used with clinical samples, failure to complete a single trial correctly or the enactment of too many incorrect moves may invoke a discontinue rule and the test is terminated (e.g. Bull et al., 2004). As the current sample include participants with healthy intellectual abilities, discontinue rules were only invoked if the participant explicitly expressed their inability to continue with a particular trial, or a desire to terminate the trials. The adoption of this liberal discontinue rule allowed for maximum performance data to be obtained across the 13-trial administration for each participant. All participants completed all 13 TOH trials without invoking the discontinue rule.

Performance on the TOH was measured by the total number of moves enacted per trial. If an error was made, the participant was alerted and asked to move the disc back to its previous position. The additional moves for this correction were added to the trial score as a penalty.

The further separation and analysis of error rates were not of primary interest to the current investigation. For each trial, the total number of moves was then subtracted from the minimum number of moves that were achievable. The resulting value was a residual score that represented how many moves beyond the minimum achievable that each participant performed.

5.2.12 Austin Maze (Darby, 2014b)

The AM is considered a measure of feedback utilisation, planning, goal setting, WM and inhibition ability (Bowden & Smith, 1994; Tucker et al., 1987). A digital iPad version (Version 2.2) of the original Milner (1965) pathway (as cited in Tucker et al., 1987) was used (Darby, 2014b). Previous research has displayed equivalency between the conventional and other digital versions of the AM (Stolwyk et al., 2013). The maze consisted of a 10x10 grid of tiles whereby participants were required to tap the tiles to learn a hidden pathway from a start to an end point whilst following a set of rules. The rules included, only being able to move one tile at a time, and only being able to move up-down, left-right and not diagonally. Participants were to complete ten trials of the maze as recommended by Bowden & Smith (1994). Scoring for the AM included the total number of error free moves (correct tiles) that were made per trial that ranged from 0 (no correct tiles) to 29 (no incorrect tiles were tapped).

5.2.13 FAS Test (Strauss et al., 2006)

Verbal fluency tests are common measures of EF and often feature amongst many neuropsychological test batteries (Lezak et al., 2012; Strauss et al., 2006). This test of phonemic fluency required participants to generate words that begin with the letters ‘F’, ‘A’, and ‘S’ with a 60-second time limit for each letter. The participant was asked to follow two rules when generating responses, (1) to not use words that are proper names, and (2) to not repeat the same word twice, even with a different suffix (e.g. ‘er’, ‘ing’). All responses were recorded, and only words that did not violate the task rules were scored as correct and awarded 1-point. In line with previous research by (Venegas & Mansur, 2011), and in an effort maximise the availability of data for analysis, the number of correct words for each 15-second interval was recorded separately. When scores were collated separate measures were included for the first 15-seconds, and then the remaining 45-seconds for each letter respectively. The FAS has previously demonstrated good internal reliability (r=.83) (Tombaugh, 1999).

5.2.14 Block Design (from the WASI) (Wechsler, 2010)

The BD was chosen for inclusion as an EF performance measure due to its ability to measure abstract conceptualisation in the context of planning (Brown et al., 2012; Lezak et al., 2012), and the requirements of the task in the context of within task environment demand changes since the number of blocks to be manipulated increases across trials. As previously explained (section 5.2.1) the traditional administration of the BD task was followed, however for the statistical analysis of within task demands the raw performance score was used for each trial (Wechsler, 2010), instead of the traditional summated score for all correct trials. As the range of raw scores does not favour a normal distribution (e.g. 0, 4, 5, 6, 7), each score was rescaled to reflect a distribution from of 1 = (incorrect design OR beyond maximum time limit), 2 (original raw score was 4), 3 (original raw score was 5) and so on, through to 5 (original raw score was 7),

The standard administration of the BD requires participants to progress through three key parameter changes. In order to replicate a design that is instructed and provided within the WASI stimulus book (Wechsler, 2010) trials 3 through 9 require participants manipulate four blocks, trials 10 and 11 require the manipulation of nine blocks, and trials 12 and 13 rotate the overall design configuration to a diamond shape (from the previous square shape in all preceding trials). The final trial, trial 13, also includes the removal of the border that specifies the overall shape of the design, which carries an additional requirement for the participant to determine whether the overall configuration of the design should be constructed as a square or diamond.

5.2.15 Re-scaling Outcome Scores

Scaling of performance across different tasks was affected by variations in scoring rules. This was most evident for tasks that use time as an outcome score where a lower result reflects better performance, compared to tasks that used summated trial scores where a higher result does the same. In an effort to ameliorate this, outcome scores were rescaled so that high values always reflected better performance, and low scores always reflected poorer performance. Where re-scaling was required, each individual participant’s score was subtracted from one numerical value above the highest recorded score for the sample. This method ensured that all characteristics of the data (e.g. variances, *SD*, distribution) remained unchanged from the original raw values.

5.2.16 Summary of Measures and Variables

A summary of scoring and the subsequent variables that were produced are presented below in Table 4.

Table 4

Summary of Measures Administered, Method of Scoring, and Corresponding Variables Produced

Measure	Scoring	Variable(s)
The Stroop Test	Time for each trial	Stroop Test– Words
		Stroop Test - Colour Words
Digit Span Backwards	Number of correct trials	Digit Span Backwards
Visual Span Backwards	Number of correct trials	Visual Span Backwards
TMT	Time for each trial	TMT-A; TMT-B
5-Point Test	Number of correct figures	5-point 0-60secs; 5-point 61-120secs;
		5-point 121-180secs
Test of d2	Number of correctly identified symbols	Test of d2
TEA- Map Search	Number of correctly identified symbols	Map Search 0-60secs;
		Map Search 61-120secs
TEA- Visual Elevator	Timing score for correct trials	Visual Elevator
TEA- Elevator Counting Reversal	Number of correct trials	ECR
TOH	Residual moves score for each trial	TOH trial 1 through to trial 13
AM	Number of correctly identified tiles	AM trial 1 through to trial 10
FAS Test	Number of correct words	F 0-15secs, A 0-15secs, S 0-15secs,
		F 16-60secs, A16- 60secs, S 16-60secs
Bock Design Subtest	Raw performance score for each trial	BD trial 1 through to trial 13

Note. TMT = Trail Making Test; ECR = Elevator Counting Reversal; TOH = Tower of Hanoi; AM = Austin Maze; BD= Block Design.

5.3 Procedure

Ethics approval was granted from the Victoria University Human Research Ethics Committee. All 103 participants were instructed to read the plain language statement (Appendix B) and if willing, completed and signed a consent form (Appendix C).

Due to the burden imposed on participants by the inclusion of wide range of neuropsychological measures, participants were given the choice of undertaking their testing session from two options; (1) 2 x 1.5 hours testing sessions, or (2) 1 x 3 hour session with a scheduled ten-minute break after 1.5 hours. Sixteen participants completed testing across two sessions, with the mean period of delay being 23.06 (*SD* = 22.68) days.

All testing was conducted in a quiet environment to minimise any external distractors at either a Victoria University campus, or at the participant's place of residence in a suitable quiet environment.

Data was collected by the author or a fellow doctoral candidate (who also used the data for a different purpose; Dr Jessica Scarfo). Both researchers were trained in neuropsychological test administration and scoring by a registered neuropsychologist (Dr Emra Suleyman). The scoring of each test was moderated between both researchers to ensure consistency and reliability of scoring. The process of triangulation was also moderated by Dr Suleyman.

5.3.1 Test administration

In an effort to minimise any confounding effects due to order, participants were randomly assigned to one of four possible test administration orders. A detailed account of counterbalancing order can be found in Appendix D. When counterbalancing consideration was given to ensure that a participant did not have any exposure a task that could have the potential to influence performance on subsequent tasks. Consideration was also given to ensure that all participants experienced the same conditions during any purposeful delay periods, and that there were no visual distractors during the delay period that could potentially influence recall. Order of presentation within tasks was strictly adhered to according to administration manuals.

The distribution of participants for each administration version, age and FSIQ-2 scores are presented in Table 5. To test whether any differences existed between administration versions on outcome and demographic variables, a series of one-way between subjects ANOVA's were conducted. A significant effect for FSIQ-2 was found at the $p<.05$ level between the 4 test administration versions, $F(3, 99) = 2.709$, $p = .049$. Post hoc comparisons using the Bonferroni test were non-significant, indicating that FSIQ-2 scores did not, in fact, significantly differ between the test administration versions. As the *F*-statistic was marginally significant, and individual pairwise comparisons returned a non-significant result, FSIQ-2 scores were considered to be comparable across test administration versions. No significant differences were found between test administration versions and participant age, $F(3, 99) = .387$, $p = .763$.

Table 5

Descriptive Statistics for Sample Size, Age, and FSIQ-2 Scores between Test Versions

Administration Version	*n*	Age		FSIQ -2 Score	
		M	*SD*	*M*	*SD*
Version A	18	112.83	9.13	29.16	7.50
Version B	22	104.90	8.25	31.18	8.06
Version C	30	107.50	11.25	30.33	6.60
Version D	33	110.90	10.41	29.36	6.88

As the statistical analysis of the current study were to utilise sample variances rather than sample means, it was important to assess whether variances were equal between the counterbalanced versions across the outcome variables of the study. As the final outcome variables of the current study were created from a composite of multiple performance scores, tests for equal variances between these variables were computed subsequent to the completion of the statistical analyses of the current study. As seen in Table 6, Levene's test of homogeneity of variances was conducted for all outcome variables, and all returned a non-significant result. Thus, the variances were not significantly different among the outcome variables across the test administration versions.

Table 6

Tests of Homogeneity of Final Outcome Variables

Outcome variable	Levene Statistic	*df*	*p*
Simple & Familiar	0.959	99	.415
Simple & Novel	1.412	99	.244
Complex & Familiar	0.891	99	.449
Complex & Novel	1.658	99	.181

5.4 Study Design

The current project was conducted across three separate studies in order to address the current aims. The overall nature of the project and each study design was iterative, with key outcomes from earlier studies being utilised during the later studies of the project. Table 7 presents an overview of each study, the key stages of analysis, and an overall purpose and/or outcome of each study.

Table 7

Overview of Each Study, Key Stages of Analysis, and Overall Purpose and/or Outcome

Study No.	Aim	Key Stages of Analysis		Purpose/Outcome
Study 1a: DCS Development and Application	Operationalise a DCS in an attempt to quantify performance according to task complexity or novelty	1a	Establish DCS framework and apply to administered neuropsychological test battery to infer the overall demand of each test environment(s).	Establish a GDC for each neuropsychological testing condition.
Study 1b: Testing of the DCS	Apply and mathematically validate the DCS to tasks that have been previously identified to produce single outcome scores, even though research has arguably demonstrated that successful performance is characterised by more than one skill.	1b	Conduct a series of Confirmatory Factor Analyses to test and explore hypothesised latent factors that represent the GDC identified during Study 1a.	Support validity of the DCS by confirming the GDC conditions for each included task element.
Study 2: Establishing GDC Models	Investigate whether test elements classified as similar in complexity and novelty hold	2.1	Create weighted composite scores for each GDC obtained in Stage 1b to enter into analysis during Stage 2.	Establish four GDC Models

Study No.	Aim		Key Stages of Analysis	Purpose/Outcome
	together mathematically to allow for comparisons not between tests, but across demands.	2.2	Conduct four one-factor congeneric measurement models to represent each Global Demand Classification using outcome scores highlighted from Studies 1a & 1b	
Study 3: Analysis of GDC	Develop a model to depict the relationships between each of the GDC Models produced in Study 2.	3.1	Use factor score regression weights obtained via *Stage 2.2* to create a single weighted composite measure of each Global Demand Classification latent construct.	A full structural model of GDC performance scores.
		3.2	Apply Hancock & Mueller's (2001) coefficient *H* formula to calculate the reliability of each composite measure.	
		3.3	Use Munck's (1979) approach to calculate factor loadings in the regression of each construct on its respective composite measure along with its associate error variance.	

Study No.	Aim	Key Stages of Analysis		Purpose/Outcome
		3.4	Specify the full structural model and perform structural equation modelling (SEM) using the single composite measure calculated during *stage 3.1* as the reflective indicator and fix both the factor loading*)* and error variances *(Step 3.3)*.	

Note. DCS = Demand Classification System; GDC = Global Demand Classification.

5.5 Statistical Assumptions Considered for Analyses conducted in Study 1b to Study 3

5.5.1 Sample Size

Determining the appropriate sample size for SEM remains a central issue within the statistical literature. Lower-bound rules of thumb often recommend *N*= 100 (Anderson & Gerbing, 1988; Ding et al., 1995; Tabachnick & Fidell, 2013). However, there is minimal consensus on the optimal value for *N* due to the variety of potential parameters and factors that may require estimation. When calculating the required *N* for an SEM study, others insist that consideration must be made for the complexity of the model, the methods of estimation used, the distribution of the data, and the degree of missingness in the data (Kline, 2015). Bentler & Chou (1987) suggested that a ratio of 5 cases per parameter is adequate for SEM when multiple indicator variables are included within a normal multivariate distribution, with no missing data. Others have recommended a ratio of 10:1 per indicator variable, particularly when deviations to normality are present (Hair et al., 2013).

Communalities have shown to have an effect on the accuracy of estimation. MacCallum et al. (1999) found that communalities of approximately .7 required a sample size of 100 to have good population recovery, with 3-4 indicator variables per latent factor. Increasing the number of variables per factor had little effect when communality was high. In light of this, the current study sample (n = 103) met the lower bound estimated for sample size assumptions of SEM. To improve the accuracy of estimation, models were not specified if they were to have less than 4-inidcator variables per latent factor (Tabachnick & Fidell, 2013).

5.5.2 Methods of Estimation

Maximum Likelihood Estimation (MLE) was chosen due to its reliability and stability during the estimation of model parameters for a Confirmatory Factor Analysis (CFA) and SEM investigations (Hair et al., 2013; Kline, 2015; Tabachnick & Fidell, 2013). An assumption of MLE is that data points are somewhat normally distributed. In the event that the distribution of the data is non-normal, but theoretically represents a true reflection of the population performance pattern (e.g. no outliers), Asymptotic Distribution-Free (ADF) estimation is often preferred as it does not assume normality within the data (Byrne, 2016; Tabachnick & Fidell, 2013). However, ADF is found to be extremely unstable in samples with <2000 cases. Thus, despite its robustness against normality during the estimation process, its application is rare amongst the social sciences (Byrne, 2016; Hair et al., 2013; Kline, 2015; Tabachnick & Fidell, 2013).

In the context of the current sample size (n=103), bootstrapped adjustments to the MLE test statistic are considered best-practice (Arbuckle, 2013; Byrne, 2016) IBM SPS AMOS 24.0 software produces the Bollen-Stine bootstrap (BS- p) which is a modification of the model chi-square statistic. The BS- p adjusts for distributions misspecification of the model due to lack of multivariate normality. Bootstrapped adjustment standard errors were also reported. In line with recommendations, all bootstrapping procedures were computed from 500-bootstrapped samples (Arbuckle, 2013).

5.5.3 Assessment of Normality

Prior to analysing the distribution of the sample data, it was important to understand potential causes of any deviations to normality that may be present within the current study. The first consideration is in reference to the characteristics of the current sample and neuropsychological test performance. Although the current sample was drawn from a non-clinical population of healthy adults, many measures within the current test battery were developed to be sensitive to performance across clinical populations where performance is typically impaired. This meant that ceiling effects were possible as non-clinical participants may be able to achieve a near maximum score, thereby causing a negative skew to some data. In the event that ceiling effects were found in high proportions, these performances scores were excluded from further analyses, due to the lack of variance that could be extracted to fulfil the assumptions of SEM analyses.

The second consideration for potential deviations to normality is the presence of outliers. From the test variables that were selected within the current study, marginal ceiling and floor effects may still have manifested for some participants. Thus, the assessment of each test variable at the univariate level also considered the nature of test performance across a multivariate distribution. For example, a person may have achieved the highest possible performance on one outcome score, but their score on another task may have fallen within the average range of performance. In order to assess this distribution, univariate and multivariate normality was analysed and reported alongside each statistical analysis conducted.

5.5.3.1 Univariate Normality

The assessment of univariate normality was carried out using a multi-step process using IBS SPSS v25.0. For each variable, skewness and kurtosis values were assessed in alignment with conventional cut-off values of -3 to 3 (Field, 2007). In the event of violations, z-score distributions were calculated for the variable. Individual cases where Z-scores fell outside the range of -3.29 to 3.29 were considered true outliers in the data. In alignment with previous approaches for the handling of univariate outliers, raw scores for outlier cases were replaced to represent a raw score of 3.29 standard deviations from the mean (Tabachnick & Fidell, 2013; Testa et al., 2012). If the variable also produced a high kurtosis value, replacing the score via this approach may have resulted in the offending case remaining an outlier (Tabachnick & Fidell, 2013). Therefore, in this instance, the case was replaced with the next largest value for that variable.

5.5.3.2 Multivariate Normality

Assessments of multivariate normality were completed using IBM SPSS AMOS 24.0. There is a common understanding that deviations in normality due to skewness tend to impact analyses of means (Byrne, 2016), whereas deviations of kurtosis can severely affect tests of variance and covariance (Byrne, 2016). As CFA and SEM are analyses of variance and covariance, highly kurtotic distributions can exert detrimental effects upon the accuracy of model estimation. As univariate distributions do not always equate to multivariate distribution, a second screening was conducted prior to interpreting each of the CFA and SEM analyses. In the context of SEM analyses and deviations from normality, kurtosis values greater than 7 are considered indicative of early departure of normality (Byrne, 2016). Mardia's normalised estimates of multivariate kurtosis were also requested. Yuan et al. (2005) suggests that Mardia's normalised estimates that are >5.00 are indicative of non-normality of the sample. However, Mardia's coefficients are reported to be unstable when samples are not very large (Byrne, 2016; Hanusz et al.). Therefore, interpretation was supplemented by assessment of Mahalanobis distance (M-distance) in light of the current sample size (n= 103).

M-distance was computed for each case to determine whether any offending cases contributed to a deviation from multivariate normality as an outlier. All maximum M-distance values were compared to a chi-squared distribution with the *df* for that analyses (Coakes, 2012). Values that exceeded the maximum chi-square threshold were suspected as an offending case. If a case was then considered a multivariate outlier, it was removed from the analysis. If no outliers were identified, but violations to multivariate normality were suspected, bootstrapped corrections to the test statistics are recommended (Byrne, 2016; Tabachnick & Fidell, 2013).

5.5.4 Model Fit Statistics

Multiple fit statistics were requested and assessed to triangulate the assessment of model fit to minimise any potential bias in the acceptance of good model fit from the assumptions of one fit statistic alone. Overall, fit indices refer to a set of calculations that provide a measure between the estimated model between the sample variance and covariance matrix and the estimated population variance and covariance matrix (Hair et al., 2013). IMB

SPSS AMOS 24.0 provided a number of fit index statistics that estimated the model fit under different assumptions (e.g. sample size, and multivariate normality) and calculations of fit parameters (e.g. *df*, residual variance).

5.5.4.1 Absolute Fit Index (x^2)

A common and well accepted measure of absolute fit for CFA and SEM analyses is the chi-square (x^2) test (Cangur & Ercan, 2015; Tabachnick & Fidell, 2013). The x^2 statistic was requested to test whether the matrix of implied variance and covariance (hypothesised model) was significantly different to the matrix of empirical sample variance and covariance (Performance data). The x^2 statistic calculates a probability level, whereby a low x^2 statistic with $p > .05$ indicates that the discrepancy between sample and parameter estimates of the model was small and representative of a good model fit. Conversely, a large x^2 statistic with $p < .05$ indicated that the differences between the sample and estimated parameters were significantly different and indicative of a poor fitting model.

5.5.4.2 Comparative Fit Indices (CFI, RMSEA)

Comparative fit indices were requested to provide a measure of the comparative fit of nested models. Within each estimated model ultimately existed a nested model that ranged from an independence model (a model that corresponds to a completely unrelated set of variables), and a saturated model (a full or perfect model) (Tabachnick & Fidell, 2013). The comparative fit index (CFI) assessed model fit relative to these other nested models. The CFI score ranges from 0 – 1, with values greater than .95 indicative of a good fitting model (Tabachnick & Fidell, 2013). However, Kline (2015) recommended that when observed parameters are <12, CFI values above .97 are favourable. The CFI has previously demonstrated robust effects when comparing model fit in smaller sample sizes (Chen, 2007; Hu & Bentler, 1999; Tabachnick & Fidell, 2013).

The Root Mean Square Error of Approximation (RMSEA) served as another well accepted estimate of comparative fit (Hair et al., 2013; Hu & Bentler, 1999; Iacobucci, 2010). The RMSEA estimated the lack of fit in the model compared to the saturated model. Values of .06 or less indicated a good fitting model, with values greater than .10 indicative of a poor fitting model (Hu & Bentler, 1999). However, Hu & Bentler (1999) recommended caution when interpreting RMSEA, as values have previously been found to over-reject true models by producing an index score above .10 when sample sizes are smaller (Hu & Bentler, 1999; Tabachnick & Fidell, 2013). Therefore, RMSEA was used as an index for comparative fit in conjunction with CFI and additional fit index scores.

5.5.4.3 Index of Proportion of Variance Accounted (GFI)

The Goodness-of-Fit Index (GFI) provided a calculation of the weighted proportion of variance in the sample covariance that was accounted for by the estimated population covariance matrix (Tabachnick & Fidell, 2013). If the proposed model is no better than an independence model, then the GFI will equal zero, however if the model provides a good fit, GFI will approach the value of 1.0 GFI values of >.95 were considered to be representative of a good fit (Hair et al., 2013; Tabachnick & Fidell, 2013).

5.5.4.4 Degree of Parsimony Fit Index (AIC)

The Akaike Information Criterion (AIC) served as a measure of parsimony within the model. The AIC is considered most useful for comparing models that are non-nested (Hair et al., 2013). When comparing AIC scores between non-nested models, the model with the smallest AIC is identified as a good fitting, and parsimonious model (Tabachnick & Fidell, 2013).

5.5.4.5 Residual-Based Fit Index (SRMR)

The assessment of model fit also required a review of standardised residual covariances to ascertain the deviation of individual covariance terms. (Hair et al., 2013) recommended that standardised residuals exceeding the range of -4.0 to 4.00 may indicate potential problems with a measurement model. Therefore, all residuals were assessed, and if considered to exceed this range, were removed from the analysis.

In addition to the assessment of individual covariances, the standardised root mean square residual (SRMR) index was calculated. SRMR provides an overall residual value that was calculated from residual differences between the sample and population variances and covariances (Tabachnick & Fidell, 2013). SRMR calculations are found to be largely independent of sample size, which reduced any bias that the conservative sample of the current study may have had on the overall assessment of model fit (Chen, 2007). General rules of thumb for a good-fitting model is an SRMR value of <.1 (hair) or alternatively .08 or less, from a possible range of 0 – 1 (Cangur & Ercan, 2015; Hu & Bentler, 1999).

A summary of fit indices and acceptable cut-off values utilises by the current project can be observed in Table 8.

Table 8

Summary of Fit Indexes and Accepted Cut-Off Values

Fit Index	Abbreviation	Acceptable Level
Chi-square	χ^2	$p > .05$
Root Mean-Square Error of Approximation	RMSEA	< .06
Standardised Root Mean-square Residual	SRMR	< .08
Goodness-of-Fit	GFI	> 0.95
Comparative Fit Index	CFI	> .95 / >.97
Akaike Information Criterion	AIC	Lowest value indicates best fitting model

5.5.5 Testing Alternative Models

Nested models exist when items from the same battery are included in a multifactorial CFA or SEM analyses. Two or more models are considered nested when one model is a subset of the other, where it is likely that strong inter-relatability exists between the constructs. For example, if a three-factor model is collapsed into a two-factor model by reducing a free parameter, the two- factor model is ultimately nested under the three-factor model. Given the analyses of many task features within this project ultimately were subsumed under an original whole global task, alternative models that eliminated any free parameters were also evaluated. Tests of discriminant validity were conducted to ascertain whether the alternative models constituted significantly better representations of the data over the hypothesised models.

Discriminant validity tests were conducted using a Nested Models method whereby free parameters were systematically collapsed by fixing their correlation to r=1 (Kline, 2015). A chi-square difference ($x^2 diff$) test was then used to test the statistical significance of the decrement in overall fit as the free parameters were eliminated, or the improvement in parameters if free parameters were added. In addition to $x^2 diff$ tests, CFI, SRMR and RMSEA indices were also informally compared, as they have previous been found to perform well in distinguishing the relative superiority of one model over another (Hair et al., 2013).

Chapter 6

Study 1a: DCS Development and Application

The strong neuroanatomical evidence for rostral-caudal activation across complex tasks, triangulated against evidence of predominantly frontal and more widespread cortical activation during EF tasks would arguably justify the convergence of neuroimaging investigations as the next phase for research investigating cognitive control during EF tasks. However, without a quantifiable framework to measure task outcomes, this endeavour would likely perpetuate the lack of clarity in both fields. Currently, cognitive control theory is limited to a theoretical account of 'task demand descriptions' without targeted objective assessment measures to test it, and EF literature is lacking robust theoretical modelling in the presence of numerous assessment measures. The imperative is therefore to develop a demand control framework to bridge these two important fields of research. The development of a framework for understanding cognitive demand should provide important insight into when cognitive control is required in relation to the specific administration, physical and mental demands that collectively establish the task environment during specific tests. Furthermore, such an approach can serve to provide comparative insight into the demands that exist between different neuropsychological tests.

Therefore, Study 1a aimed to establish a Demand Classification System (DCS) to provide a framework for structured appraisal and classification of demands for complexity and novelty of the task environment during the performance of goal-directed adaptive behaviour using neuropsychological tests of EF. Like many other researchers and seminal theories (e.g. Anderson, 2002; Normal & Shallice, 1986: Stuss & Benson, 1984), this project will consider attention as a foundational component of higher-order EF performance, and as such measures of attention will be included in the consideration of task demands.

The DCS is designed as an evaluation framework founded on the overarching guiding concepts of complexity and novelty. As such, demands are conceptualised across a dual-axis framework, where each axis represents a continuum of complexity and novelty respectively. Each of these two axes is then applied to the test environment of the selected neuropsychological/ EF measures.

The conceptualisation of a dual axis framework for the DCS gives rise to four hypothesised Global Demands, with the continuum of complexity spanning from (S) Simple to (C) Complex, and the continuum of novelty spanning from (F) Familiar to (N) Novel. Along each axis, demand is conceptualised as increasing distinguishably. At the lower extremes of these axes, (S) Simple, and (F) Familiar demands are considered to represent a task environment requiring little to no recruitment of cognitive control, and at the upper extremes (C) Complex, and (N) Novel demands represent a task environment requiring recruitment of high levels of cognitive control.

A set of Demand Criteria for assessment and quantification of the four Global Demands (S, C, F and N) was developed based on previous research that has demonstrated the conditional and environmental demand features that call for recruitment of cognitive control, as outlined in Chapter 1. Specifically, the complexity axis is assessed across (1) Abstraction, (2) Contextual Stability, (3) Action Rules, (4) Instructions and Rules, and (5) Dual Nature. As (1) Abstraction is regarded as an umbrella term for three types of Abstraction demands, three specifiers were assigned, including (T) Temporal Abstraction, (P) Policy Abstraction, and (R) Relational Abstraction.

The Novelty axis is assessed across (6) Automaticity, (7) Schematic Demands, and (8) Episodic Demands. The DCS, including each Demand Criterion, is presented in Table 9.

Table 9

The Demand Classification System (DCS)

ID	Demand Features		Demand Criteria		
			(S) Simple		(C) Complex
(1)	Abstraction	1.C		2.C	
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response
(2)	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue
(3)	Action Rules	3.S	A minimal subset of possible response options are available	3.C	Many alternative response options are available
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment
(5)	Dual Nature	5.S	Task requires completion of a singular task	5.C	Task encompasses a dual or multi-task requirement
			(F) Familiar		(N) Novel
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements

Each of the Demand Criteria were given an identifier that comprised of a number corresponding to the relevant Demand Feature for each Global Demand, together with a letter corresponding to the specific Demand Criterion (S, C, F or N). For Example, Simple Action Rules was assigned the identifier (3.S), whereas Complex Action Rules was assigned the identifier (3.C).

6.1 Method

6.1.1 Procedure

Prior to DCS application, each test was appraised as to whether it should be considered (a) Singular: the entire task condition is largely unchanging, and thus the traditional outcomes scores were considered adequately reflective of demand or, (b) Multifaceted: with different items needing to be grouped and scored due to dynamic changes in their requirements. All tests appraised as multifaceted were previously identified in the literature as containing natural divisions in task performance. For multifaceted tests, consideration needed to be applied to the conceptualisation of their scoring systems. Where a multifaceted test provided an existing scoring system consistent with differences in task demands, this was used. In the absence of an existing scoring system that encapsulated multifaceted performance, two different approaches were taken. Firstly, if insufficient evidence was available from the literature to allow clear demarcation of differences based on demand, a qualitative appraisal of the test administration environment was made for any natural changes in administrative requirements and expectations on performance. The divisions resulting from this appraisal were triangulated with two Psychologists. Finally, if previous literature has suggested that divisions exist, but the structure of these divisions could not be clearly identified (e.g. Tower paradigm), exploratory analyses were performed to determine if divisions could be elicited based on test performance.

6.1.2 DCS Appraisal and Scoring

The characteristics of each test were appraised in relation to administration and physical demands, responses and actions necessary for successful performance outcome. Demand Criteria were the exemplar from which specific features/characteristics were identified, such as, stable vs. changing stimuli; the addition/removal of stimuli; the need to evaluate, re-configure, manipulate, or repeat stimuli; the need to follow, adapt or appraise any endogenous and exogenous signals that would guide a response. This appraisal is highlighted by the following example:

A task issues a series of 10 different size blocks that are required to be assembled into a configuration that matches an exogenously prescribed design. The task consists of five trials, with each trial presenting a new design that must be constructed. The number of blocks provided do not change in size or quantity between each trial. In parallel to this task, the participant is also required to accurately recite a story that was presented prior to the start of the block task.

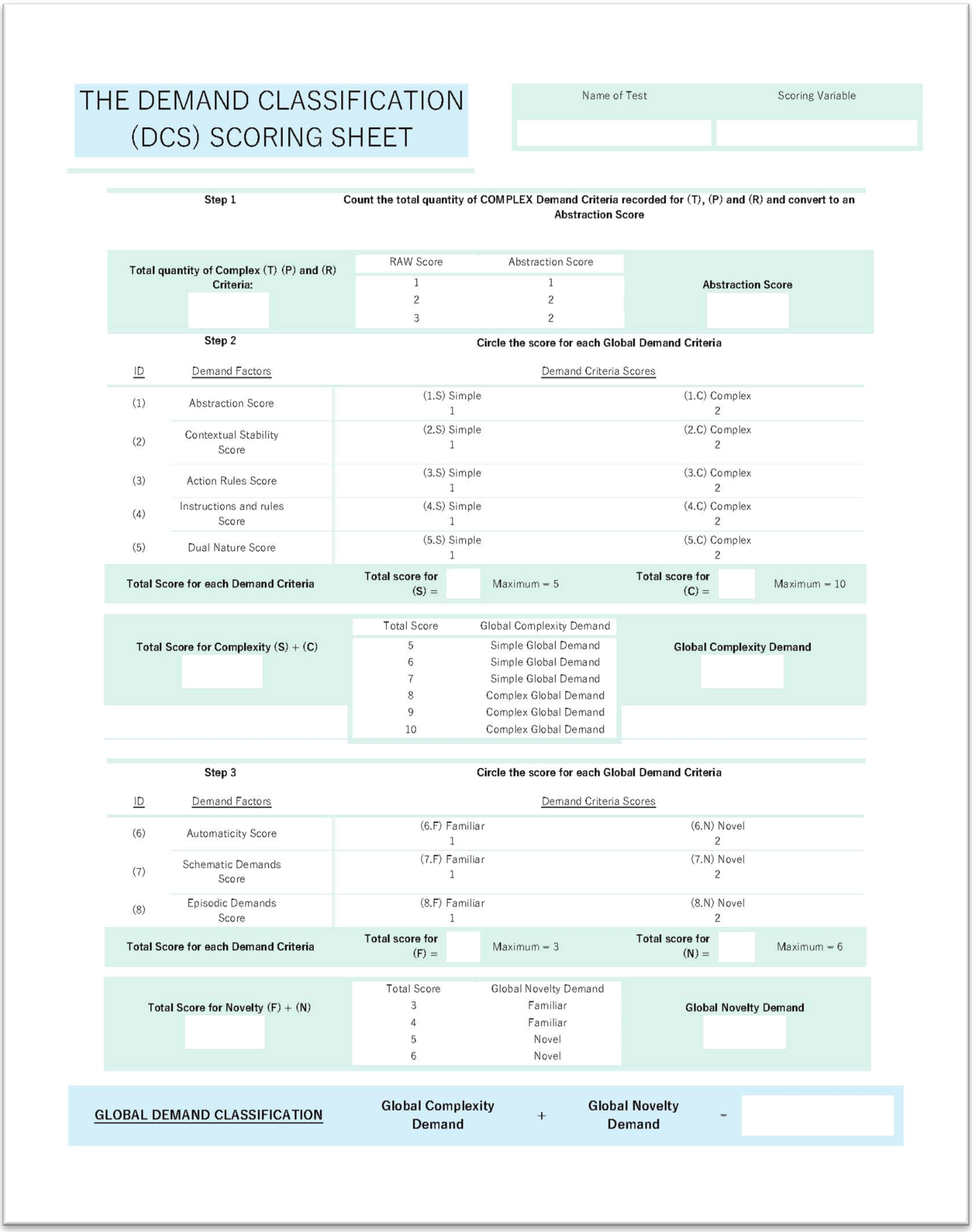
When appraised against the Demand Criteria, this task would be classified to encompass (1R.C) Complex Relational Abstraction Demands due the assessment of relationships that is required between the different sizes of blocks and how they can be utilised as a whole to construct the prescribed design. The task would also be classified to encompass a (5.C) Complex Dual Nature, due to the individual being required to complete two separate tasks simultaneously. Furthermore, the task would also be classified to encompass (8.F) Familiar Episodic Demands due to each trial being independent of the other, as performance is not dependent on the retention of the design configuration across all trials.

In order to determine the overall GDC of each neuropsychological test, a DCS Record Sheet was created (Figure 1). The outcome for each Demand Criteria were entered into the scoring sheet. This was then used to determine separate scores for the Global Demands of

Complexity and Novelty, which were further used to determine an overall GDC score for each task.

Figure 1

The Demand Classification (DCS) Record Sheet.



THE DEMAND CLASSIFICATION (DCS) SCORING SHEET

Name of Test	Scoring Variable

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Criteria:	RAW Score	Abstraction Score	Abstraction Score
	1	1	
	2	2	
	3	2	

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** Maximum = 5	**Total score for (C) =** Maximum = 10

Total Score for Complexity (S) + (C)	Total Score	Global Complexity Demand	Global Complexity Demand
	5	Simple Global Demand	
	6	Simple Global Demand	
	7	Simple Global Demand	
	8	Complex Global Demand	
	9	Complex Global Demand	
	10	Complex Global Demand	

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** Maximum = 3	**Total score for (N) =** Maximum = 6

Total Score for Novelty (F) + (N)	Total Score	Global Novelty Demand	Global Novelty Demand
	3	Familiar	
	4	Familiar	
	5	Novel	
	6	Novel	

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** =

Note. The scoring sheet is used in conjunction with the DCS Criteria (Table 9).

6.1.2.1 Simple and Complex Demand Criteria

In order to distinguish scores along the continuum from (S) Simple to (C) Complex, each of the (C) Complex Demand Criteria were allocated a score of =2, and each of the (S) Simple Demand Criteria were allocated a score of =1. Equal weighting was assigned within each Global Demand Criteria given that each (C) Complex Demand Criteria was representative of a particular environmental demand that has previously been found to require cognitive control. It remains inconclusive as to whether the demands for any one of the three specifiers of abstraction (temporal, policy or relational) are more invoking of cognitive control than another. Moreover, one form of abstraction is rarely found to feature alone within testing paradigms (Badre & Nee, 2018). Therefore, all three abstraction criteria served as specifiers to inform the overall Global Demand for (1) Abstraction. As seen in Figure 1, any two or more abstraction specifiers that were found to encompass two or more (S) Simple or (C) Complex demands were considered to represent overall (S) Simple or (S) Complex Demands for Abstraction.

6.1.2.2 Familiar and Novel Demand Criteria

In order to distinguish scores along the continuum from familiar to novel, each of the (N) Novel Demand Criteria were allocated a score of =2, and each of the (F) Familiar Demand Criteria were allocated a score of =1. Equal weighting was also assigned to each Demand Criteria given that each (N) Novel Demand Criteria was representative of an environmental demand that has previously been found to require cognitive control.

6.1.2.3 Calculating an Overall DCS Score

For a test to be classified to encompass (C) Complex Global Demands, a total score of ≥8 was required, whereas a total score of ≤7 classified the test to encompass (S) Simple Global Demands. For a test to be classified to encompass (N) Novel Demands, a total score of ≥5 was required, whereas a total score of ≤4 classified the test to encompass (F) Familiar Demands.

The scoring of the DCS was designed as an accumulative record system and not a definitive additive system. An additive system would convey the assumption that all Demand Criteria within the DCS are independent and that the relative complexity within and between each is known. This would require the contribution of each Demand Criteria to be known in order to assign a numerical representative weighting. The implementation of an additive system was considered inappropriate within the DCS due to the interdependent nature of the testing environment that is represented by the Demand Criteria. An accumulative system was considered most appropriate as this approach acknowledges that all Demand Criteria contribute collectively towards an overall GDC, rather than the emphasis on the value of the impact of any single Demand Criteria over another, as is the case with an additive approach (Modave & Shokar, 2013). The accumulative scoring system of the DCS was considered well-suited for enabling post hoc evaluation and comparisons between the four DCS Global Demands. The DCS is not intended to be used to directly infer the comparative complexity and novelty of task items within each GDC.

In consideration of the aforementioned, the GDC represented two of four possible Global Demands for a test environment. The combinations of Global Demands could be classified as either Simple & Familiar (S&F), Simple & Novel (S&N), Complex & Familiar (C&F), or Complex & Novel (C&N). These GDCs captured each quadrant across the dual axis, as an appraisal could not yield a score of =0 on either continuum. The outcome of the DCS for each test under appraisal was considered a hypothesised GDC that required further analyses during Studies 1b, 2, and 3 of the project. The final hypothesised GDC of the DCS were considered to be representative of the continuum of complexity and novelty of either (a) a whole test environment, or (b) a set of given trials/tasks within a whole test environment that were found to encompass different Global Demand.

In consolidation of the aforementioned inner mechanics of the DCS, the DCS application first yields a Global Demand (S, N, F, or N) for each demand axis (complexity and novelty). A Global Demand Classification is then calculated based on this outcome which reconciles these demands across a dual axis; S&F, S&N, C&F, or C&N.

6.2 Results

Of the 13 Neuropsychological tests included in the test battery, five tests were classified as singular (Table 10). These included Digit Span Backwards, Visual Span Backwards, Test of d2, Visual elevator, and Elevator Counting with Reversal.

Table 10

Summary of Singular Test/Task Scoring and Requirements

Test/Task Element(s)	Scoring	Description
Digit Span Backwards	Total number of correct trials	Requires verbatim repetition of numbers in reverse order measured via the small addition (+1) of similar parameters (digits) for every two correct trials
Visual Span Backwards	Total number of correct trials	Require repetition of ordered sequencing of stimuli in reverse measured via the small addition (+1) of similar parameters (blocks) for every two correct trials.
Visual Elevator	Total timing score for correct trials	Incremental changes to the quantity of elevator stimuli and the directionality of counting within the same paradigm
Elevator Counting Reversal.	Total number of correct trials	Incremental changes to the quantity of elevator stimuli and the directionality of counting within the same paradigm
Test of d2	Total number of correctly identified configurations	Requires continued search for target letter configuration

Note. Each test/task element scores are total scores.

Four tests were classified as multifaceted and demonstrated published scoring that was consistent with the natural divisions within the task. These included the Stroop Test, TMT, 5-Point test, and Map Search. All these tests scores were included as per standardised published instructions (outlined in Section 5.2) and the features within them are provided in Table 11.

Table 11

Summary of Multifaceted Task Elements, Scoring, and Requirements that were Consistent with Natural Demand Divisions.

Test	Task Element(s)	Scoring	Description
Map Search	Map Search 0-60secs	Total number correctly identified stimuli	Requires the search and identification of target stimuli
	Map Search 61-120secs	Total number correctly identified stimuli	Continued search and identification of target stimuli
5-Point Test	5-point 0-60secs	Total number of correct figures	Initial drawing of unique figures
	5-point 61-120secs	Total number of correct figures	Continued drawing of unique figures
TMT	TMT A	Total Time	Drawing of a continuous line between numbers in numerical order spread across a page
	TMT B	Total Time	Drawing a continuous line alternating between numbers and letters in numerical and alphabetical order
The Stroop Test	Stroop Test – Words	Total time	Reading aloud series of word colour names
	Stroop Test – Colour-Word	Total time	Reading aloud series of colour names printed in a mismatched coloured ink

Note. TMT= Trail Making Test.

Three tests were classified as multifaceted (Table 12), with clear divisions apparent at specific junctures within the test (BD, AM, FAS Test), and one final test warranted further classification of the task demands within and was subjected to further exploratory statistical analysis (The TOH). These tests are outlined further below.

Table 12

Summary of Multifaceted Task Elements, Scoring, and Requirements where Divisions were Apparent at Specific Junctures.

Test	Task Element(s)	Scoring	Description
BD	Trials 5, 6, 7, 8 and 9	Raw performance score	Requires the manipulation a *four* block square design
	Trials 10 and 11	Raw performance score	Requires the manipulation a *nine* block square design
	Trials 12 & 13	Raw performance score	Requires the manipulation a *nine* block *diamond*, with Trial 13 seeing the removal of the border that specifies the overall shape if of the design
FAS Test	F' 0-15secs, 'S' 0-15secs, 'F' 15- 60secs	Total number of correct words	Initial generation of verbal vocabulary
	'F' 16- 60secs, 'A' 16-60secs, 'S' 16-60secs	Total number of correct words	Continued production of verbal vocabulary
AM	Trials 7, 8 & 9	Number of correctly identified tiles	Requires the immediate execution and recall of the newly presented hidden maze
	Trial 3, 4, & 5	Number of correctly identified tiles	Requires the recall of the learned hidden path

Note. BD= Block Design; AM= Austin Maze.

As indicated in Table 12, from the original 13-trials of the BD Test, nine trials were included for further analysis in the current study. Trials 1 & 2 were excluded from analysis as the reverse rule was not invoked within the current sample and no performance data collected. Trials 3 & 4 were also excluded due to these being learning trials, and their standard scoring criteria does embed a measure of time to complete each of these trials.

Trials were also excluded from the original 10 AM trials. Trial 1 was omitted from further analysis due to its execution being entirely exploratory. During Trial 1, the participant has no knowledge or exposure to the hidden path and begins to find the path via trial-and-error. Trial 10 was also omitted as 52.3% of the current sample had reached a < 2-error trials at this stage, demonstrating that the majority of the sample were approaching the optimal performance score. Whilst performance on Trial 10 provides clinical importance at the individual level of assessment to infer learning of the maze, within the current data set the high proportion of the sample that obtained this score reflect the ceiling effects of later AM trials. Furthermore, Trials 2 and 6 were considered to promote ambiguity within the current appraisal. Performance during Trial-2 often requires the continuation of an exploratory search pattern, and Trial-6 may represent the mid-way point of the task, as 55.3% of the participants at this stage had reached < 4 error-free trials. Thus, the removal from analysis was considered appropriate to enable a more direct measure of the study aims. Further analysis was conducted during later stages of this project (study 1b) to ascertain the weather exclusion of these trials was appropriate.

6.2.1 Exploration of the Tower of Hanoi (TOH)

The TOH was selected for appraisal due to previous literature suggesting that different demands exist within its configuration. Kaller et al. (2011) devised a software program to assist in defining the problem structure underlying performance on Tower tasks. The TOH employed for the current project (Bishop et al., 2001) was entered into the TowerTool v2.0 software to enable an in-depth analysis of its problem structure (Kaller et al., 2011). As seen in Table 13, the problem structure within the TOH varied across many parameters. Similarities between trials could be found under any one TOH parameter, but no clear similarities could be identified when all parameters were considered. Without clear evidence of the relative influence that each parameter can have over TOH complexity, the generation of a priori theoretical appraisal was not feasible.

Table 13

The Internal Problem-Structure of the TOH

	TOH Parameters					
	Configuration					
Trial	Start	Goal	Min Moves	Disks	Counterintuitive Moves	Sub-goals
3	Tower	Flat	4	3	0	3
4	Flat	Tower	5	3	2	3
5	Tower	Flat	5	3	1	3
6	Flat	Tower	6	3	2	3
7	Tower	Flat	6	3	1	3
8	Tower	Tower	7	3	2	3
9	Tower	Flat	7	3	3	3
10	Flat	Tower	8	4	2	4
11	Tower	Flat	8	4	2	4
12	Flat	Tower	9	4	3	3
13	Tower	Flat	9	4	2	3

Note. TOH= Tower of Hanoi.

As such, further exploratory analysis of TOH performance was required in order to ascertain whether any trials shared statistical communality. Prior to this analysis TOH trials 1 & 2 were excluded as their capacity to provide a reflective measure of performance in a complex goal directed task was considered minimal. TOH trials 1 & 2 served as learning trials that were used to ascertain the participant's understanding and comprehension of the task and its rule set. During these trials, participants could ask questions, and apply the task rules to a simple and small problem structure. Thus, individual performance effects of the TOH were not considered to have the ability to manifest until Trial 3.

6.2.1.1 Exploratory Factor Analysis of TOH

An Exploratory Factor Analysis (EFA) using MLE with an oblique rotation (direct oblimin) was executed using 11 trials of the TOH on data from 103 participants to assess its underlying factor structure prior to entry into a CFA. MLE was considered the most appropriate extraction method due to its ability to extract the common unique shared variance, which is considered a preferential approach when identification of latent constructs is desired (Hair et al., 2013). An oblique rotation was applied as it was expected that TOH trials would correlate given that they are subsumed under the same global task.

6.2.1.1.1 Assumptions

Assumptions of multivariate normality and linearity were evaluated using IBM SPSS v25.0. As shown in Table 14, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality. No significant outliers were identified via analyses of M-distance. Mardia's estimate (Table 14) suggested potential violations for multivariate normality. Whilst EFA is considered robust against deviations from normality, BS-*p* was requested during the post hoc CFA to adjust for any potential distributional misspecification. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy statistic was within an acceptable range to infer that the data is suitable for factor analysis (.571), and Bartlett's Test of Sphericity was significant (<.001) supporting the assumption that TOH variables were significantly different for an identity matrix. There were no missing data.

Table 14

Descriptive Statistics for TOH variables

Variable	*N*	*M*	*SD*	Range		Skewness	Kurtosis
				Potential	Actual		
TOH 3	103	4.34	4.36	1- ∞	1.00-18.00	1.437	1.113
TOH 4	103	1.86	2.36	1- ∞	1.00-10.00	2.699	5.857
TOH 5	103	4.74	4.17	1- ∞	1.00-18.00	-0.955	0.034
TOH 6	103	2.71	2.45	1- ∞	1.00-10.00	1.364	0.957
TOH 7	103	4.20	3.95	1- ∞	1.00-17.00	1.275	0.714
TOH 8	103	3.25	3.30	1- ∞	1.00-13.00	1.316	1.885
TOH 9	103	3.28	2.94	1- ∞	1.00-13.00	1.760	2.825
TOH 10	103	8.57	11.72	1- ∞	1.00-48.00	1.938	2.923
TOH 11	103	11.44	13.46	1- ∞	1.00-53.00	1.727	2.296
TOH 12	103	6.41	8.19	1- ∞	1.00-31.00	1.922	2.755
TOH 13	103	10.50	12.19	1- ∞	1.00-51.00	1.570	2.043
Mardias Estimate							6.660

Note. TOH= Tower of Hanoi, ∞ = Uncapped maximum score.

6.2.1.1.2 Factor Structure of the TOH

Four factors with Eigenvalues exceeding >1 were identified as underlying the 11-item TOH paradigm. In total, these factors accounted for 54% of the variance in the TOH data. When an oblique rotation was requested, correlations between all four factors were low (Appendix E). Consequently, an orthogonal rotation (Varimax) was then chosen. Under conventional recommendations, three or more items loading onto one factor is considered a minimum requirement to infer stability and accuracy (Hair et al., 2013; Tabachnick & Fidell, 2013). As seen in Table 15, only Factor 1 met this criterion. Therefore, the trials that were represented by Factor 1 were considered the only trials to represent a task element suitable for appraisal by the DCS.

Table 15

Varimax Rotated Factor Structure of the 11-Item TOH

Item	Factor Loadings			
	Factor 1	Factor 2	Factor 3	Factor 4
Trial 3	**0.412**	-0.036	-0.040	0.190
Trial 5	**0.653**	0.102	0.028	-0.229
Trial 7	**0.385**	0.044	-0.139	0.299
Trial 8	**0.620**	-0.043	0.038	0.080
Trial 11	**0.308**	-0.026	-0.006	0.074
Trial 9	-0.018	**0.367**	-0.140	0.113
Trial 12	-0.014	**0.988**	0.148	0.022
Trial 10	-0.149	-0.161	**0.963**	0.157
Trial 13	0.160	0.034	-0.172	**0.487**
Trial 4	0.038	0.027	0.246	-0.114
Trial 6	0.029	0.069	0.040	0.299

Note. Numbers in bold type face represent the significant loadings for each factor.

6.2.1.1.3 Tower of Hanoi Problem Structure Analysis

TOH trials 3, 5, 7, 8 and 11 were then re-assessed using TowerTool v2.0 software. As seen in Table 16, a shared parameter of all TOH trials was a *tower* starting configuration. Whereas, goal configurations were *flat-ending* for the majority of trials, with the exception of Trial 8. Each TOH trial had a different number of minimum moves. The number of *subgoals* that were required were similar for all 3 -disk trials, with an additional *subgoal* required for the 4-disk Trial 11. Furthermore, with the exception of Trial 3, all trials required the enactment of counterintuitive moves.

Table 16

TOH Problem Structure Parameters

Trial	Configuration		Min Moves	Disks	Counterintuitive Moves	Sub-goals
	Start	Goal				
Trial 3	Tower	Flat	4	3	0	3
Trial 5	Tower	Flat	5	3	1	3
Trial 7	Tower	Flat	6	3	1	3
Trial 8	Tower	Tower	7	3	2	3
Trial 11	Tower	Flat	8	4	2	4

6.2.2 Outcome of DCS Application per Singular or Multifaceted Test.

A total of 20 GDCs were demarcated by the DCS application. The following Table 17 summarises outcomes of the DCS appraisal for each task element, their scoring variables, their DCS Criteria and final GDC. Completed scoring sheets and DCS Criteria for each Task Element can be viewed in Appendix F.

Table 17

Summary of DCS Criteria and Global Demand Classifications for included Test Elements.

	Demand Criteria										
Task Element	(1) Abstraction	(2) Contextual Stability	(3) Action Rules	(4) Instructions and rules	(5) Dual Nature	(6) Automaticity	(7) Schematic Demands	(8) Episodic Demands	Complexity Score	Novelty score	Global Demand Classification
Stroop Test – Words	S	S	S	S	S	F	F	F	5	3	S&F
Block Design Trials 5, 6, 7, 8, & 9	C	S	C	S	S	F	F	F	7	3	S&F
TMT-A	S	S	S	S	S	F	F	F	5	3	S&F
FAS Test 0-15secs	S	S	C	S	S	N	F	N	6	4	S&F
5-Point Test 0-60secs	S	S	C	C	S	F	F	F	7	4	S&F
Map Search 0-60secs	S	S	S	S	S	F	F	F	5	3	S&F
Test of d2	S	S	S	S	S	N	F	F	5	4	S&F
Visual Span Backwards	C	S	C	S	S	N	N	N	7	6	S&N
Digit Span Backwards	C	S	S	S	S	N	N	N	5	6	S&N
FAS Test ‘F’ 16- 60secs, ‘A’ 16- 60secs, & ‘S’ 16-60secs	S	S	C	S	S	N	F	N	6	5	S&N
5-Point Test 61-120secs	S	S	C	C	S	N	F	N	7	5	S&N
Visual Elevator	S	S	S	C	S	N	N	F	6	6	S&N
Block Design Trials 10 & 11	C	C	C	S	S	F	F	F	8	4	C&F
Map Search 61-120secs	C	C	C	S	S	F	F	F	8	3	C&F
Austin Maze Trials 7, 8 & 9	C	S	C	C	S	N	F	F	8	4	C&F
TMT-B	C	C	C	S	C	N	F	F	8	4	C&F
TOH 3, 5, 7, 8, & 11	C	C	C	C	S	N	F	F	9	4	C&F
Bock Design Trials 12 & 13	C	C	C	S	S	N	N	N	8	6	C&N
Stroop Test Colour-Word	C	C	S	C	S	N	N	F	8	5	C&N
Austin Maze Trial 3, 4, & 5	C	S	C	C	S	N	F	N	8	6	C&N
Elevator Counting Reversal	C	S	C	C	C	N	N	F	9	5	C&N

Note. S&F = Simple & Familiar; S&N = Simple & Novel; C&F = Complex & Familiar; C&N = Complex & Novel

6.3 Discussion

Study 1a aimed to establish the DCS to assist with the identification of demands for cognitive control that may exist across a battery of neuropsychological tests. Complex and novel environmental features are often targeted when assessing the efficiency of cognitive control of an individual. To do this, task paradigms are administered that establish a testing environment where objective novelty and complexity feature. However, an empirical approach to the identification and evaluation of these demands remains absent from the literature. Therefore, the DCS sought to identify the demand for cognitive control by evaluating task-specific elements against criterion for complexity and novelty.

6.3.1 Simple and Complex Global Demands

The structure of the DCS allowed for a contrasting appraisal of (S) Simple vs. (C) Complex demands of a testing environment. Overall, this appraisal considered whether a task afforded direct engagement with singular unchanging stimuli, or the need to overcome ambiguity, potentially across multiple stimuli. The appraisal of these demands across five different Demand Features provided a detailed insight into how the environment of a task can vary and supported the demarcation of demands based on the administrative requirements of the testing environment.

Task elements were identified that represented conditions whereby minimal cognitive control resources may be required. Within the current test battery, 11 tasks were identified to reach a (S) Simple Global Demand score, with four tests (Stroop – Words; TMT-A, Map Search 0-60secs, Test of d2) found to be comprised of exclusively (S) Simple demand criteria in relation to the axis representing the continuum of Complexity. Collectively, these task environments represent the need for direct engagement with a singular stimulus, where explicit S→R relationships are required to be followed. Given the availability of exogenous instructions that guided the response during these task elements, performance under (S) Simple Global

Demands are considered to require minimal cognitive control resources for completion (Brass et al., 2017; Longman et al., 2019; Verbruggen et al., 2018; Wenke et al., 2015).

In contrast, complex testing environments were identified to predominantly reflect conditions where ambiguity exists between the S→R. The nature of these test environments is likely to elicit the support of cognitive control due to the need to deliberately generate strategies in order to reduce ambiguous relationships within the tasks. Overall, (C) Complex task demand environments were found to reduce performance due to purposefully unclear parameters to establish a direct response to a given stimulus. Previous research has proposed the engagement of cognitive control during these conditions arise from the additional management of abstract contexts (Badre & D'Esposito, 2007), cognitive tracking of multiple items, and the integration of multiple information sources (Nee et al., 2014). Interestingly, no single tests were found to fulfil all (C) Complex Demand Criteria, however all tasks that were classified under (C) Complex Global Demands encompassed (1.C) Complex Abstraction Demands. As Abstraction can be considered a central hallmark to the engagement of cognitive control in complex environments (Badre & Nee, 2018), the ability to identify these task conditions via the DCS provides further insight into where abstraction abilities may be best captured during test performance.

A prominent finding was that the majority of multifaceted test environments were identified to comprise of both (S) Simple and (C) Complex Global Demands. The ability to capture and identify this variability supports the current position of this thesis that demands within test environments must be appraised across multiple Demand Features. This approach respects previous research that proposed cognitive control engagement to be responsive to the gradient of complexity and is not modular by nature (Badre & D'Esposito, 2007; Hugdahl et al., 2015; Jeon & Friederici, 2015). The ability to identify the specific Demand Features within each task serves to enforce this approach and offer insight into the diversity of demands that are placed on the individual during any given test performance. Furthermore, the representation of these Demand Features using the DCS criteria of (S) Simple and (C) Complex demands provides insight into how task demands may be offset by each other. For example, the Stroop Colour-Word task element was identified to have (C) Complex Demands due to ambiguity between S→R relationships, changing context and unique task-specific rules. However, this demand may be offset by the small number of action rules are offered (Steinbeis & Crone, 2016; Vohs et al., 2008), and the singular response nature of the task (Olszanowski & Szostak, 2019).

6.3.2 Familiar and Novel Global Demands

The appraisal of demands for novelty within a testing environment allowed for a contrasting evaluation between tasks that enabled the application of fundamental knowledge, and those that require the formation of new knowledge. Testing environments were found to be exclusively (F) Familiar when their execution required application of automatic responses and implicit skillsets only. This was identified amongst tasks that required reading (Stroop Words), counting (TMT-A) and searching abilities (Test of d2, Map Search). Performance within these tasks are likely supported by long-term memory retrieval processes that enable automaticity of cognitive processing, and therefore an overall reduction in cognitive control effort (Wheatley & Wegner, 2001).

Conversely, (N) Novel test environments were identified that required the uptake of a new series of actions, and/or the flexibility to adapt implicit knowledge or schemas to new task information. The (N) Novel Demand Criteria mostly reflected test environments that require cognitive control due to the need for rational and deliberate thought (Badre & Nee, 2018). This was represented by the Digit Span-Backwards and Visual Span- Backwards Tasks where the retention of novel episodic information is required to form a new sequence of behaviour.

Moreover, these tasks require the counteracting of well-learned automatic S→R due to an opposing instructed or rule-governed action of the reverse order rule.

Importantly, within multifaceted tests, the DCS was able to distinguish testing environments where the performance outcomes are contingent on the ability to learn from initial novel stimuli (Austin Maze). This allowed the identification and demarcation of task elements that represent the novel components of the task, versus the task elements where the novelty of the task is expected to become learned. During the initial novel task environment, the AM bares unfamiliar S→R features. Verbruggen et al. (2014) proposed that during these demand conditions, cognitive control can be considered largely reactive in its activation. In contrast, during the familiar environment of the task, cognitive control has the capacity to be proactive given the initial experience that allowing for the preparation and self-regulation of the anticipated S→R (Verbruggen et al., 2014). The identification of these varied demands for novelty is suggestive of a varied level of demand for cognitive control within the singular test environment, which potentially represents two different performance outcomes not offered by traditional scoring.

The onset of novel demands within a familiar environment can also cause cognitive resources to be disrupted from any ongoing task performance (Barcelo et al., 2006). BD was found to introduce novel demands during its later trials (BD Trials 11 & 12). This novel demand arises due to an increase in schema acquisition and integration of previous S→R experience to adapt to the new BD design configuration. These novel requirements provide a key indicator to *where* recruitment of cognitive control may increase within a test environment (Wirzberger et al., 2018).

Historically, novelty has proven difficult to infer amongst neuropsychological tests due to it necessitating the absence of previous experience with the test environment. This challenge may in part be due to many testing environments comprising of both familiar and novel demands, which the DCS has attempted to capture.

6.3.3 The Global Demand Classification (GDC)

The GDC was designed to acknowledge that demands can occur across dual axes of complexity and novelty. As conceptualisation of intersecting axes implies, application of the DCS to the current test battery identified that complexity and novelty rarely occur in isolation, and instead test environments were identified where complexity and novelty simultaneously occurred at varying degrees. Therefore, considering demands across a single axis alone would fail to identify where cognitive control resources may be required. For example, the Digit Span–Backwards test was identified to encompass (S) Simple demands due to a singular context, instructed environment, and linear increase in difficulty related to capacity. However, the task was also identified to encompass (N) Novel demands due to the need to recall digits in a reverse order. Thus, failure to consider the novelty within the test would have misrepresented its demands for cognitive control resources. Similarly, tests were found to comprise of (F) Familiar demands which if considered on a singular axis, would wrongly assume that minimal cognitive control is required. Therefore, the use of a singular axis conceptualisation may have wrongly called into question the use of some tests for assessment of cognitive control or EF. Within the current battery, this could have resulted in 11 testing environments to be considered as encompassing minimal demands for cognitive control, including the much-used TMT-B (Lezak et al., 2012). However, by considering the presence of complexity alongside demands for novelty across a dual axis, the appropriate recognition of the cognitive control demands of each task was facilitated.

Appraising task demands across a dual axis also identified a selection of task environments where (C) Complex demands could be influenced by the coexistence of a (F) Familiar demand environment. For example, during C&F test environments, the identification of complex ambiguous relationships between stimuli can be supported by implicit knowledge and familiarity of the testing environment. When environmental demands allow for a continuous direct interaction with this declarative knowledge of the environment, schemas can be accessed, and the appropriate response formulated with minimal cognitive control effort. This was highlighted within Badre & Nee's (2018) research, where the overall demand for abstraction is considered to depend on the structure of schemas held in memory stores. Thus, while cognitive control is likely still required during C&F demands, this overall demand is reduced due to the ability to access and apply already known contextual knowledge and skillsets.

By utilising a GDC, the overall demands for cognitive control are able to be appropriately represented as an interaction of (C) Complex and (N) Novel demand features, with C&N GDC being reflective of a task environment that is most requiring of cognitive control resources. This approach coalesces with previous literature that has proposed a gradient of response to demand by cognitive control (Badre & D'Esposito, 2007; Badre & Nee, 2018; Duncan, 2013; Jeon & Friederici, 2015; Koechlin & Summerfield, 2007; Pisula et al., 2019). Thus, the ability to classify test environments across a dual axis of demand allows for this gradient of demand for cognitive control to become operationalised. In doing so, comparisons can be made between the GDC that exist across an entire test battery.

6.3.4 Conclusion

Study 1a has provided an account of demands across a continuum of complexity and novelty for cognitive control in relation to the Demand Features within a test environment. This account enabled both the identification and preliminary evaluation of similarities and differences in these demand features both *within* and *between* different neuropsychological tests. The calculation of a GDC for each task element provides a representation of where demands may exist across a dual axis of complexity and novelty. The provision of this understanding offers an initial benchmark whereby individual differences may be gauged and contrasted in respect to demand performance. Further investigation is required to ascertain whether performance outcomes within these testing environments are influenced by the demands that have been identified via the DCS. This investigation must explore whether performance actually reflects each GDC to determine the competency of the DCS to identify demand conditions. The offering of a framework that is able to represent demand performance may provide an empirical alternative towards understanding test performance within the neuropsychological testing environment.

Chapter 7

Study 1b: Testing of the DCS

The aim of Study 1b was to mathematically validate the DCS for tasks traditionally represented by a singular outcome score, even though research has arguably demonstrated that successful performance is underpinned by more than one skill. The application of the DCS during Study 1a demonstrated that varied Global Demands can be identified within multifaceted tests of EF. Thus, the central premise of Study 1b was to ascertain whether performance outcomes within these testing environments are influenced by the demands that have been identified using the DCS. Hypotheses will not be explicitly stated, but follow the application of the DCS as described in Study 1a and are representative figuratively during Study1b. The outcomes from this study will subsequently inform the creation of a weighted composite variables during later chapters of the overall project.

7.1 Method

Study 1a provided the theoretical foundation for a series of CFAs where each GDC was used to inform the structure of a representative latent factor model for each test and their indicator variables. The tests under investigation were the BD, AM, FAS Test, and TOH since application of the DCS in Study 1a implied varying GDC not captured by their traditional scoring systems. GDC models were testing using performance scores obtained from each hypothesised test element.

Statistical analyses were performed using IMB SPSS AMOS v25.0, and data analysis was performed in conjunction with statistical design and conventions outlined in Sections 5.5 and 5.6.

When latent factors are discussed within the text, they are denoted by capitalisation. Any latent factors that demonstrated both statistical and theoretical significance were given the descriptor of a 'GDC Factor' in relation to Global Demand Classification for each test element under investigation.

For any accepted models, t-tests were performed between the mean scores for each GDC Factor. This served to inform the nature of performance under each GDC Factor by providing further insight into whether performance scores differed significantly from each other.

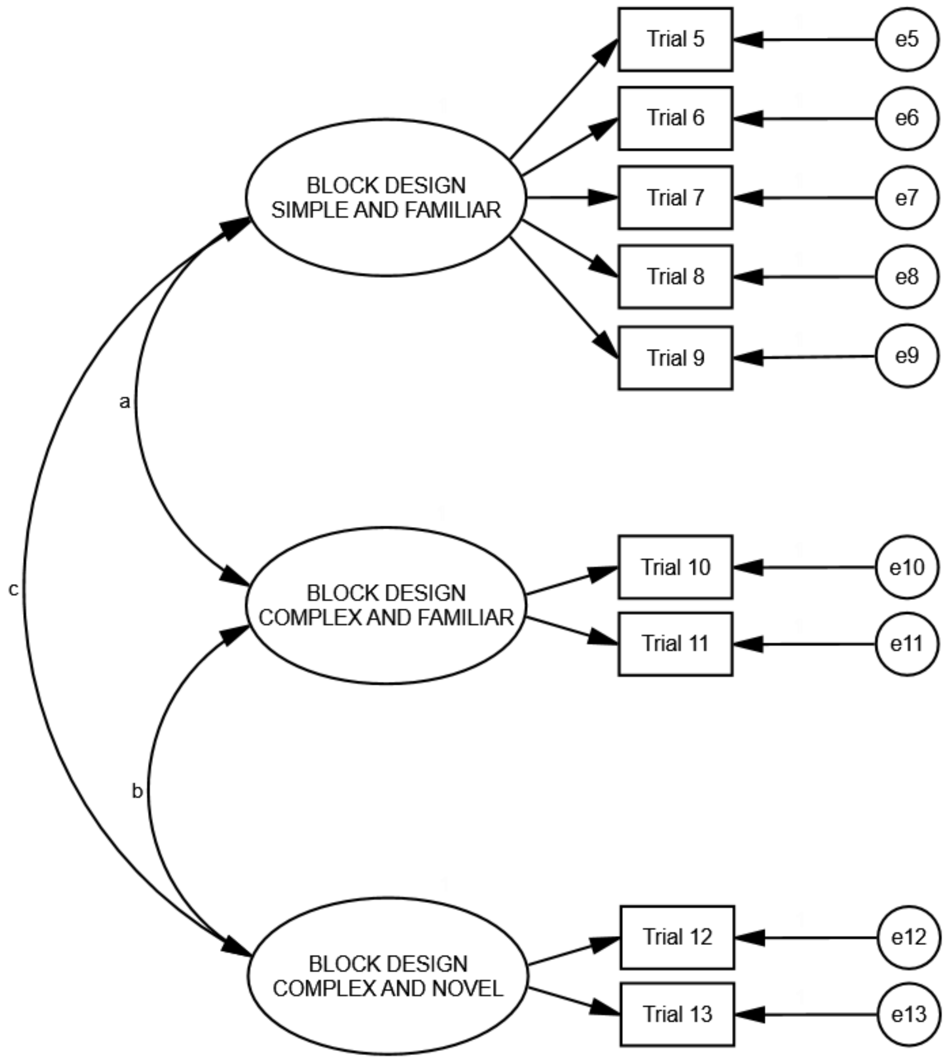
7.2 Results

7.2.1 Block Design

A CFA on nine trials from the BD task was performed. The hypothesised model in Figure 2 examined whether three GDCs were representative of performance variance within the BD Task. As the three factors were ultimately nested within a one common factor, alternative models were also specified by collapsing two factors into one by constraining their correlation to $r = 1$ (Figure 2).

Figure 2

Hypothesised Model of Block Design Factor Structure



Note. Letters a, b, and c denote the fixed correlation value r = 1 between latent factors for model comparison. The larger circles represent endogenous latent variables, squares represent exogenous indicator variables, and the small circles represent the residual variance.

7.2.1.1 Assumptions

As shown in Table 18, Skewness and Kurtosis values fell within the acceptable range to support the assumption of univariate normality. M-distance did not exceed the critical x^2 for *df* = 24 (α=.001) of 51.18 for any cases, indicating that multivariate outliers were not of concern. Mardia's estimate (Table 18) was within the acceptable range to support the assumption of multivariate normality. There were no missing data.

Table 18

Descriptive Statistics for Block Design Indicator Variables

				Range			
Variable	*N*	*M*	*SD*	Potential Score	Actual Score	Skewness	Kurtosis
Trial 5	103	3.18	1.16	1- 5	1.00-5.00	0.229	-1.269
Trial 6	103	3.64	0.93	1- 5	2.00-5.00	-0.305	-0.759
Trial 7	103	4.14	1.06	1- 5	2.00-5.00	-0.986	-0.344
Trial 8	103	3.22	1.14	1- 5	1.00-5.00	0.067	-1.036
Trial 9	103	3.41	1.24	1- 5	1.00-5.00	-0.251	-1.117
Trial 10	103	3.81	1.23	1- 5	1.00-5.00	-0.870	-0.272
Trial 11	103	2.88	1.25	1- 5	1.00-5.00	0.041	-1.048
Trial 12	103	2.54	1.25	1- 5	1.00-5.00	0.248	-1.078
Trial 13	103	1.98	1.98	1- 5	1.00-5.00	0.989	-0.134
Mardia's estimate							-.264
M- Distance							20.395

7.2.1.2 Model Estimation and Comparison

A CFA using MLE was performed using data from 103 participants. As shown in Table 19, the three factor model demonstrated a good fit between the model and observed data x^2(24, *103*) = 17.202, *p*= .840, SRMR = 0.0326, RMSEA = <.001, CFI =1.00, GFI = .963. No aberrant residuals were reported, thus supporting the fit of the model (Appendix G).

As shown in Table 19, each of the three alternative two-factor models also returned a non-significant x^2 statistic, indicating the appropriate fit of all alternative models. From the alternative models, a trend in the fit statistic values indicated that the model of best fit occurred when the BLOCK DESIGN COMPLEX & FAMILIAR and BLOCK DESIGN COMPLEX & NOVEL latent factors were collapsed into one single latent factor (Table 19).

To accurately assess whether the alternative two-factor model provided significant improvement to model fit, a nested models x^{2-}difference test was computed. As shown in Table 19, marginal improvement was seen when BLOCK DESIGN SIMPLE & FAMILIAR & BLOCK DESIGN COMPLEX & FAMILIAR latent factors were collapsed into one latent factor, $\chi^2_{diff}(1, n = 103) = 3.550, p = .060$ in comparison to the greater improvements in model fit indices (Table 19) when the BLOCK DESIGN COMPLEX & FAMILIAR and BLOCK DESIGN COMPLEX & NOVEL factors were collapsed into one single latent factor $\chi^2_{diff}(1, n = 103) = .269, p = .681$.

Table 19

Model Fit Statistics and Comparisons of Block Design GDC Model

	Model Fit								Fit vs. Full Model		
Model	*df*	χ^2	*p*	SRMR	RMSEA	GFI	CFI	AIC	χ^2_{diff}	*df*	*p*
Full three factor model	24	17.202	.840	.0326	<.001	.963	1	--	--	--	--
Two- factor models											
(a) Factor 1: BD SIMPLE & FAMILIAR & BD COMPLEX & FAMILIAR constrained. Factor 2: BD Complex & Novel Factor	25	20.752	.706	.0366	<.001	.955	1	60.752	3.550	1	.060
(b) Factor 1: BD COMPLEX AND FAMILIAR & BD COMPLEX & NOVEL constrained. Factor 2: BD SIMPLE & FAMILIAR	**25**	**17.371**	**.868**	**.0328**	**<.001**	**.963**	**1**	**57.371**	**0.269**	**1**	**.681**
(c) Factor 1: BD SIMPLE & FAMILIAR & BD COMPLEX & NOVEL constrained. Factor 2: BD COMPLEX & FAMILIAR	25	23.121	.571	.0387	<.001	.952	1	63.121	5.919	1	.015
One-factor model	27	29.037	.359	.0447	.027	.936	.994	--	11.835	3	.008

Note. The Model that indicated best fit is displayed in bold type face. BD = Block Design; SRMR= Standardised Root Mean Square Residual; RMSEA = Root Mean-Square Error of Approximation; GFI Goodness of Fit; Comparative Fit Index; AIC = Akaike Information Criterion. SRMR <.08, RMSEA <.05, GFI >.95, and CFI >.95 indicate good fit. AIC with the lowest value indicates best fit for non-nested models. Lower-case letters in parentheses preceding each two factor model comparisons represent the relationships depicted in Figure 2.

7.2.1.3 Model Re-specification

To ensure parsimony of the model fit, the model was re-specified into a two-latent factor solution. In addition, a revision of the DCS criteria for each collection of trials representative of each latent factor was required in the context of a two latent factor model. Application of the DCS to the trials relevant to this solution resulted in the initial criteria for the BLOCK DESIGN SIMPLE & FAMILIAR being reconceptualised as more representative of a C&F GDC (Appendix H). This classification was considered most suited to the complexity of the block utilisation and design structure requiring use of both full coloured and split-half coloured sides. The change to this criterion meant that the latent factor that represented Trials 5-9 was considered C&F. The trials subsumed under the collapsed COMPLEX latent factors (which removed hypothesised notions of difference between trials 10, 11 and 12, 13) were considered to encompass C&N demands within the BD task (Appendix H). The final GDC Factors, along with the BD trials they represent are presented in Table 20.

Table 20

GDC Factors Established via Block Design CFA

GDC Factor	Indicator Variables
BD C&F	Trial 5
	Trial 6
	Trial 7
	Trial 8
	Trial 9
BD C&N	Trial 10
	Trial 11
	Trial 12
	Trial 13

Note. BD = Block Design; C&F= Complex & Familiar; C&N= Complex & Novel.

As shown in in Table 21, MLE of the two-factor BD model demonstrated significant relationships between all BD trials and their corresponding latent factors.

Table 21

Maximum Likelihood Estimates for CFA of Block Design Trials

	Estimates				
Parameters	Standardised	Unstandardised	SE	Residual[a]	*p*
Trial 5 ← BD C&F	.692	0.805	.107	.706	< .001
Trial 6 ← BD C&F	.597	0.557	.090	.561	< .001
Trial 7 ← BD C&F	.681	0.719	.098	.598	< .001
Trial 8 ← BD C&F	.726	0.823	.103	.614	< .001
Trial 9 ← BD C&F	.728	0.905	.113	.725	< .001
Trial 10 ← BD C&N	.710	0.872	.112	.750	< .001
Trial 11 ← BD C&N	.750	0.937	.112	.681	< .001
Trial 12 ← BD C&N	.813	1.012	.108	.525	< .001
Trial 13 ← BD C&N	.662	0.783	.110	.785	< .001

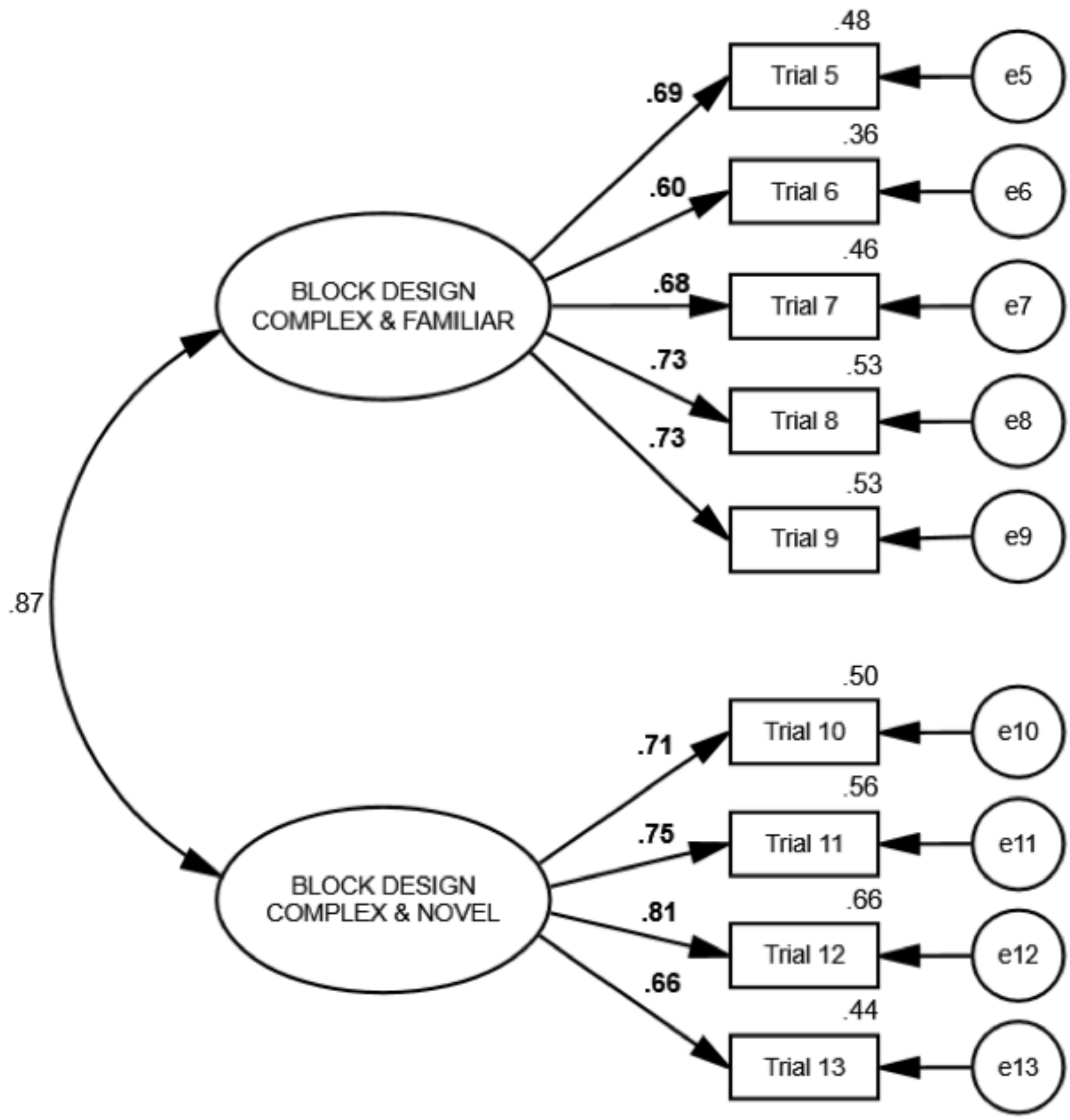
Note. SE = Standard Error; BD = Block Design; C&F = Complex and Familiar; C&N Complex and Novel.
[a]Unstandardised error variance associated with each trial.

A paired sample t-test was conducted to compare mean performance scores between BD C&F and BD C&N GDC Factors. There was a significant difference in the mean scores for BD C&F (M= 3.52, SD= .84) and BD C&N (M= 2.80, SD= .996) GDC Factors; $t(102)$= 10.166, p= <.001. The final model, including coefficients in their standardized form, is illustrated in Figure 3.

Figure 3

Final Model from CFA of Block Design GDC Factor Structure

.



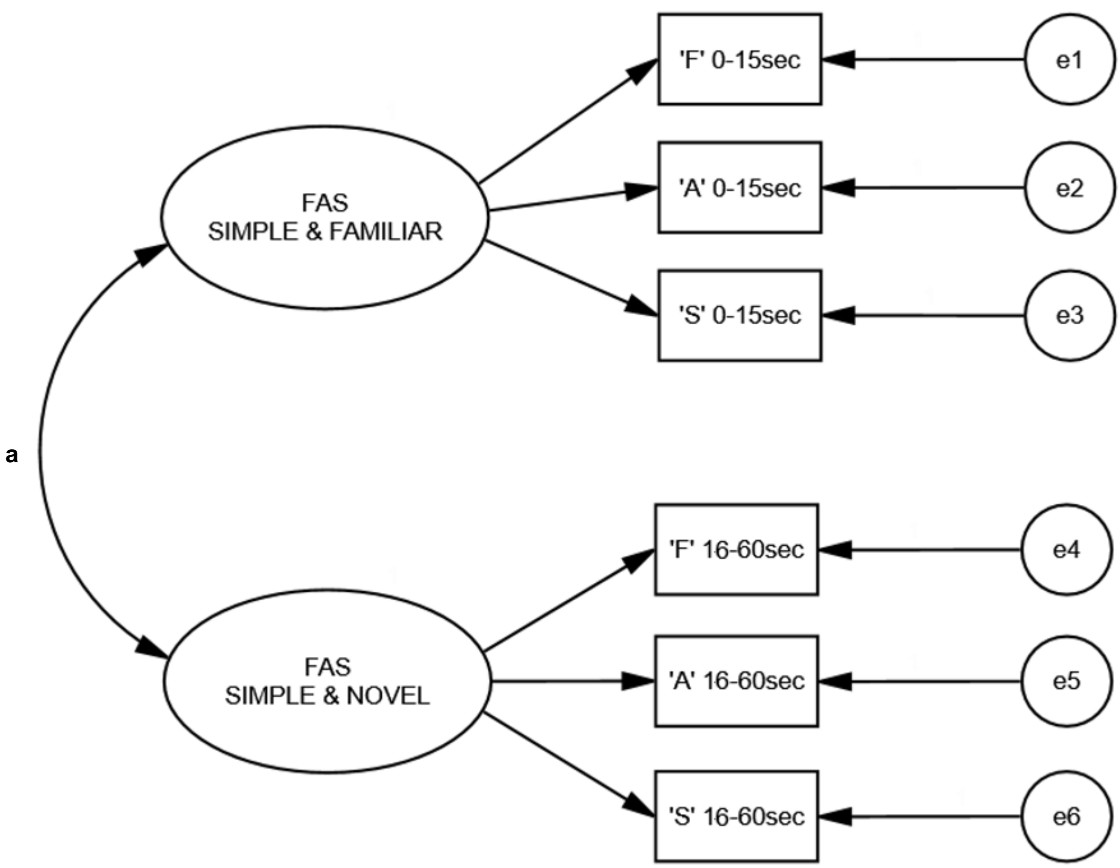
Note. The number adjacent to the curved double-headed arrow is the inter-factor correlation. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each exogenous variable are the squared multiple correlations.

7.2.2 FAS Test

A CFA was performed on two trials of the FAS test. The hypothesised model shown in Figure 4 examined whether two GDCs identified by the DCS underpinned performance across the duration of the FAS test.

Figure 4

Hypothesised Model of FAS Test GDC Factor Structure



Note. Letter ‘a’ denotes the fixed correlation value $r = 1$ between latent factors for post hoc model comparison. The larger circles represent endogenous latent variables, squares represent exogenous indicator variables, and the small circles represent the residual variance.

7.2.2.1 Assumptions

As shown in Table 21, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality. M-distance did not exceed the critical x^2 for $df = 8$ ($\alpha = .001$) of 26.12 for any cases, indicating that multivariate outliers were not of concern. Mardia's estimate (Table 22) was within the accepted range to assume multivariate normality. There were no missing data.

Table 22

Descriptive Statistics of FAS Indicator Variables

				Range			
Variable	*N*	*M*	*SD*	Potential Score	Actual Score	Skewness	Kurtosis
'F' 0-15sec	103	6.18	2.19	0 - ∞	1.00 -13.00	0.303	0.207
'A' 0-15sec	103	4.48	1.69	0 - ∞	1.00 -10.00	0.508	0.415
'S' 0-15sec	103	5.81	1.80	0 - ∞	3.00 -10.00	0.348	-0.525
'F' 16-60sec	103	2.70	1.12	0 - ∞	0 - 5.33	0.155	-0.159
'A' 16- 60sec	103	2.33	1.05	0 - ∞	0.33-5.00	0.660	-0.187
'S' 16-60sec	103	3.43	1.20	0 - ∞	1.00-6.00	0.244	-0.784
Mardia's estimate							-1.046
M-distance							13.61

Note. ∞ = No maximum score is set by the task.

7.2.2.2 Model Estimation and Comparison

A CFA using MLE was performed using data from 103 participants. As shown in Table 23, the hypothesised two-factor model demonstrated a good fit across all indices, $x^2(8, \mathit{103}) = 10.866$, $p = .209$, SRMR = .043, RMSEA = .059, CFI = .997, GFI = .964. An evaluation of the standardised residuals did not reveal any aberrant values that would indicate poor fit (Appendix I). The hypothesised model also demonstrated a significantly superior fit to the data in comparison to a nested one factor model as constraining model parameters to a one-factor model significantly worsened model fit, χ^2 diff (1, n=103) = 8.433, p= .004.

Table 23

Model Fit Statistics and Comparisons of FAS GDC Model

	Model Fit							Fit vs. Hypothesised Model		
Model	*df*	χ^2	*p*	SRMR	RMSEA	GFI	CFI	χ^2 diff	*df*	*p*
Hypothesised Model	**8**	**10.866**	**.209**	**.043**	**.059**	**.964**	**.997**			
One factor Model	9	19.299	.023	.062	.106	.941	.922	8.433	1	.004

Note. The model that indicated best overall fit is displayed in bold type face. SRMR= Standardised Root Mean Square Residual; RMSEA = Root Mean-Square Error of Approximation; GFI Goodness of Fit; CFI= Comparative Fit Index; SRMR <.08, RMSEA <.05, GFI >.95, and CFI >.95 indicate good fit.

As shown in in Table 24, MLE of the hypothesised two-factor model demonstrated significant relationships between all indicator variables and their corresponding latent factors. SMC values indicated that a good proportion of variance in the indicator variables was explained by each of the latent factors, with the exception of the *'S' 0-15sec* indicator (SMC= .213). Notwithstanding the low amount of explained variance of the *S' 0-15sec* indicator, the indicator was still a statically significant estimate. By following statistical convention, the indicator was retained in the model. The retention of the *'S' 0-15sec* indicator was also theoretically justifiable given its embedded nature within a continuing task condition, which provided a good measure of the FAS S&N factor.

Table 24

Maximum Likelihood Estimates from CFA of FAS Test

Parameters	Estimates		SE	Residual[a]	*p*
	Standardised	Unstandardised			
'F' 0-15sec ← FAS S&F	.580	1.270	.241	3.178	<.001
'A' 0-15sec ← FAS S&F	.727	1.223	.188	1.336	<.001
'S' 0-15sec ← FAS S&F	.461	0.828	.200	2.534	<.001
'F' 16- 60sec ← FAS S&N	.697	0.778	.111	0.639	<.001
'A' 16- 60sec ← FAS S&N	.722	0.760	.104	0.530	<.001
'S' 16-60sec ← FAS S&N	.699	0.835	.119	0.730	<.001

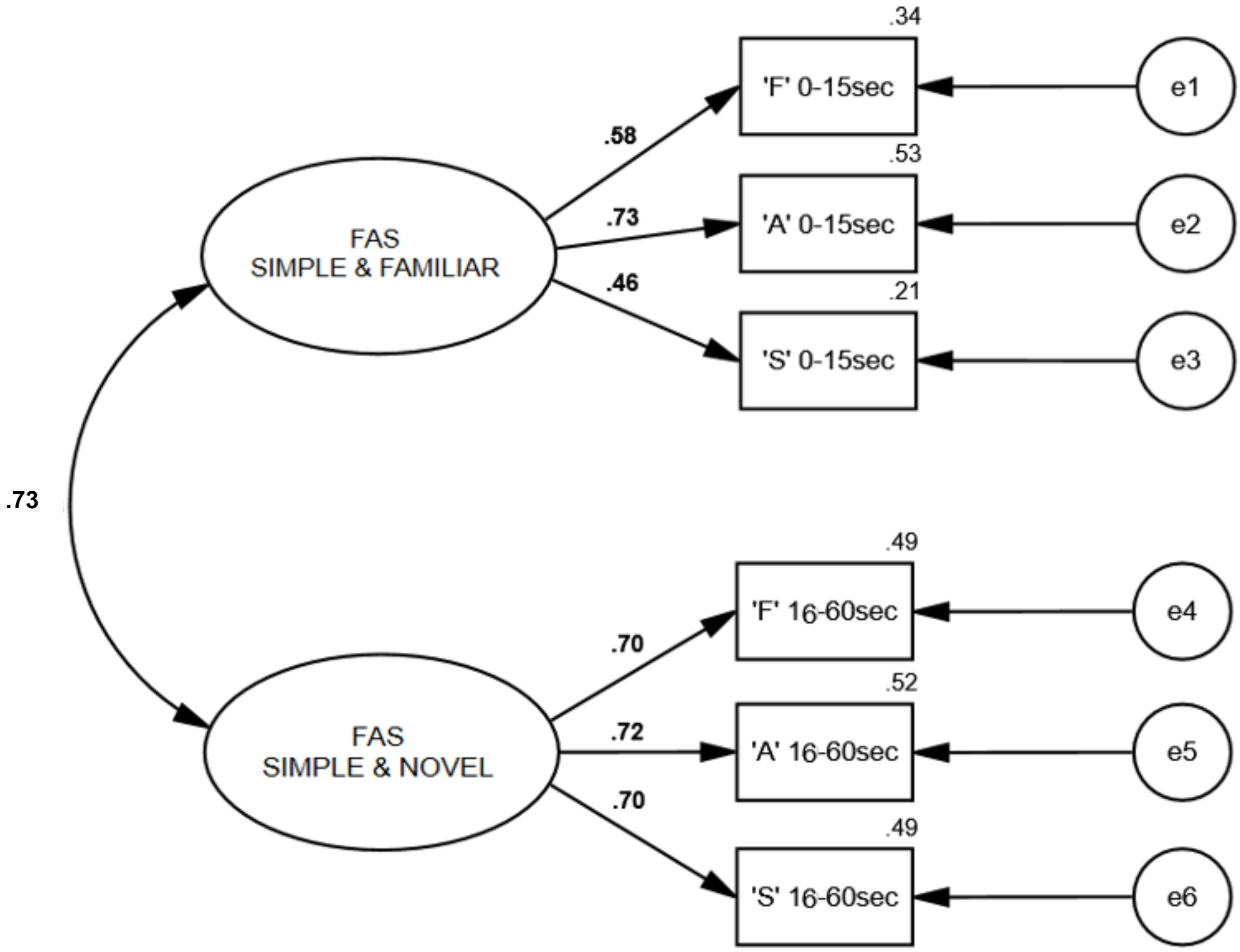
Note. SE = Standard Error; S&F = Simple and Familiar; S&N = Simple & Novel.
[a]Unstandardised residual error variance associated with each trial.

A paired sample t-test was conducted to compare mean performance scores between FAS S&F and FAS S&N GDC Factors. There was a significant difference in the mean score for FAS S&F (*M*= 5.49, *SD*= 1.421) and FAS S&N (*M*= 2.82, *SD*= .918) GDC Factors;

$t(102)=21.70$, $p= <.001$. The final model, including coefficients in their standardized form, is illustrated in Figure 5.

Figure 5

Final Model from CFA of FAS GDC Factor Structure



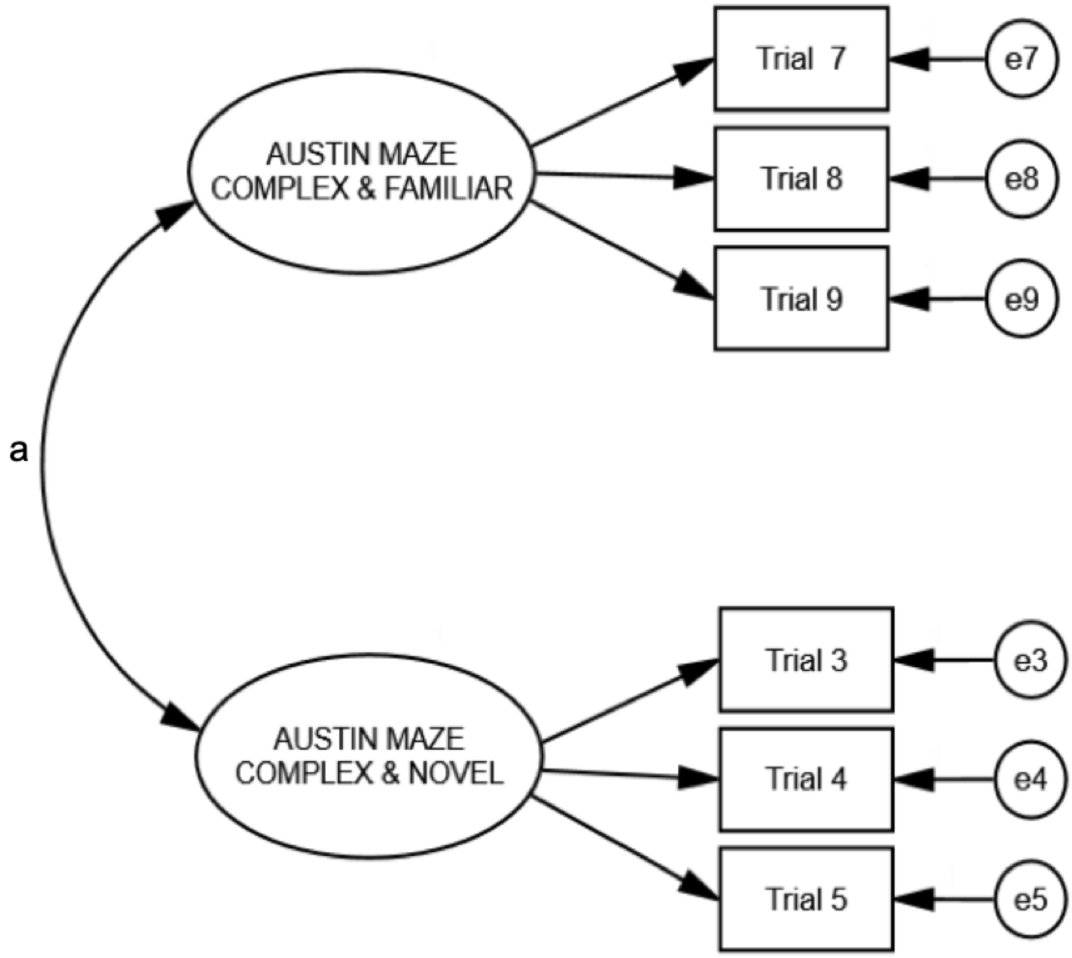
Note. The number adjacent to the curved double-headed arrow is the inter-factor correlation. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each exogenous variable are the squared multiple correlations (SMC).

7.2.3 Austin Maze

A CFA was performed to investigate the GDC of the AM task. It was unclear which latent factor Trial-2 and Trial-6 would best represent. The hypothesised two-factor model (*M0*) presented in Figure 6 included six of the ten trials, with Trials 2 and 6 excluded. In an effort to ensure the most reliable reflective indicators of each hypothesised factors were included from the AM, alternative model comparisons were also carried out that included additional free parameters for Trials 2 and Trials 6. *M1* included the additional of Trial 6 to the COMPLEX & FAMILIAR latent factor. M2 included the addition of Trial 2 to the COMPLEX & NOVEL latent factor.

Figure 6

Hypothesised Model of Austin Maze GDC Factor Structure



Note. Hypothesised model for Austin Maze factor structure. Letter ‘a’ denotes the fixed correlation value $r = 1$ between latent factors for post hoc model comparison. The larger circles represent endogenous latent variables, squares represent exogenous indicator variables, and the small circles represent the residual variance.

7.2.3.1 Assumptions

As shown in Table 25, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality. Mardia's estimate signified the potential violation of multivariate normality. M-distance did not exceed the critical x^{2} for df = 19 (α =.001) of 43.82 for any cases, indicating that multivariate outliers were not of concern. Due to suspected violations to multivariate normality, BS-*p*, and bootstrapped standard errors were requested to adjust for any distributional misspecification of the model. There were no missing data.

Table 25

Descriptive Statistics for the Austin Maze

Variable	*N*	*M*	*SD*	Range		Skewness	Kurtosis
				Potential	Actual		
Trial 2	103	17.44	4.63	1-28	3 - 26	-1.104	-1.606
Trial 3	103	19.97	4.36	1-28	6 - 27	-1.156	1.333
Trial 4	103	21.45	3.59	1-28	10 - 27	-1.099	0.957
Trial 5	103	22.60	3.47	1-28	11 – 28	-1.187	1.460
Trial 6	103	23.36	3.28	1-28	13 – 28	-1.022	0.799
Trial 7	103	23.99	3.60	1-28	12 – 28	-1.316	1.413
Trial 8	103	24.62	3.07	1-28	15 – 28	-1.131	1.290
Trial 9	103	24.83	3.07	1-28	15 - 28	-1.315	1.290
Mardia's estimate							8.807
M-Distance							33.414

7.2.3.2 Model Estimation and Comparison

A CFA using MLE was performed using data from 103 participants. As shown in Table 26, *M0* demonstrated an good fit between the model and observed data $x^2(8, \mathit{103}) = 9.445$, *p*= .306, BS-*p* = .423, SRMR = .223, RMSEA = .042 , CFI = .997, GFI = .974. *M0* also demonstrated a significantly superior fit to the data in comparison to a nested one factor model χ^2_{diff} (1, *n* = 103) = .38.222, *p* = .<001. An evaluation of the standardised residuals did not reveal any aberrant values that would indicate poor fit (Appendix J).

Alternative models *M1* and *M2* were also analysed. As shown in Table 26, both alternative models returned a non-significant χ^2, demonstrating a good fit to the data. For *M1*, an *SMC* of 0.325 was recorded for Trial 2, demonstrating a low reliability in comparison to the other trials (Appendix J). A χ^2_{diff} test indicated a significant improvement in model fit for *M1* in comparison to *M0* χ^2_{diff} (11, *n* = 103) = .18.422, *p* = .<072. However, as shown in Table X, RMSEA and GFI values fell outside the threshold for acceptable model fit, and SRMR and CFI indices worsened in comparison to *M0* fit indices. *M2* also demonstrated poorer overall fit across all fit indices in comparison to *M0* (Table 26), and significantly worsened the model fit χ^2_{diff} (5, *n* = 103) = .12.726, *p* = .<026. In consideration of the accumulation of fit indices for all models analysed, *M1* and *M2* resulted in a poorer overall fit in comparison to model *M0*, and were therefore rejected.

Table 26

Model Fit Statistics and Comparisons of Austin Maze GDC Model

	Model Fit								Fit vs. Hypothesised Model		
Model	*df*	χ^2	*p*	SRMR	RMSEA	GFI	CFI	BS-*p*	χ^2 diff	*df*	*p*
Hypothesised Model (*M0*)	**8**	**9.445**	**.306**	**.022**	**.042**	**.974**	**.997**	**.423**	**-**	**-**	**-**
One factor Model	9	47.767	<.001	-	.205	.615	.925	.005	38.222	1	<.001
Alternative Model 1 (*M1*)	19	27.867	.086	.032	.068	.938	.987	.303	18.422	11	.072
Alternative Model 2 (*M2*)	13	22.171	.053	.028	.083	.943	.986	.209	12.726	5	.026

Note. The model that indicated best overall fit is displayed in bold type face. SRMR= Standardised Root Mean Square Residual; RMSEA = Root Mean-Square Error of Approximation; GFI= Goodness of Fit; CFI= Comparative Fit Index. SRMR <0.08, RMSEA <.05, GFI >.95, and CFI >.95 indicate good fit.

As shown in in Table 27, MLE of the two-factor *M0* model demonstrated significant relationships between all indicator variables and their corresponding latent factors.

Table 27

Maximum Likelihood Estimates for Confirmatory Factor Analysis of Austin Maze

	Estimates				
Parameters	Standardised	Unstandardised	SE[a]	Residual[b]	*p*
Trial 7 ← AM C&F	.909	3.262	.371	2.241	< .001
Trial 8 ← AM C&F	.873	2.669	.305	2.219	< .001
Trial 9 ← AM C&F	..868	2.653	.338	2.304	< .001
Trial 3 ← AM C&N	.824	3.578	.472	6.060	< .001
Trial 4 ← AM C&N	.898	3.215	.326	2.494	< .001
Trial 5 ← AM C&N	.908	3.139	.371	2.096	< .001

Note. SE = Standard Error; AM = Austin Maze; C&F = Complex and Familiar; C&N = Complex and Novel.

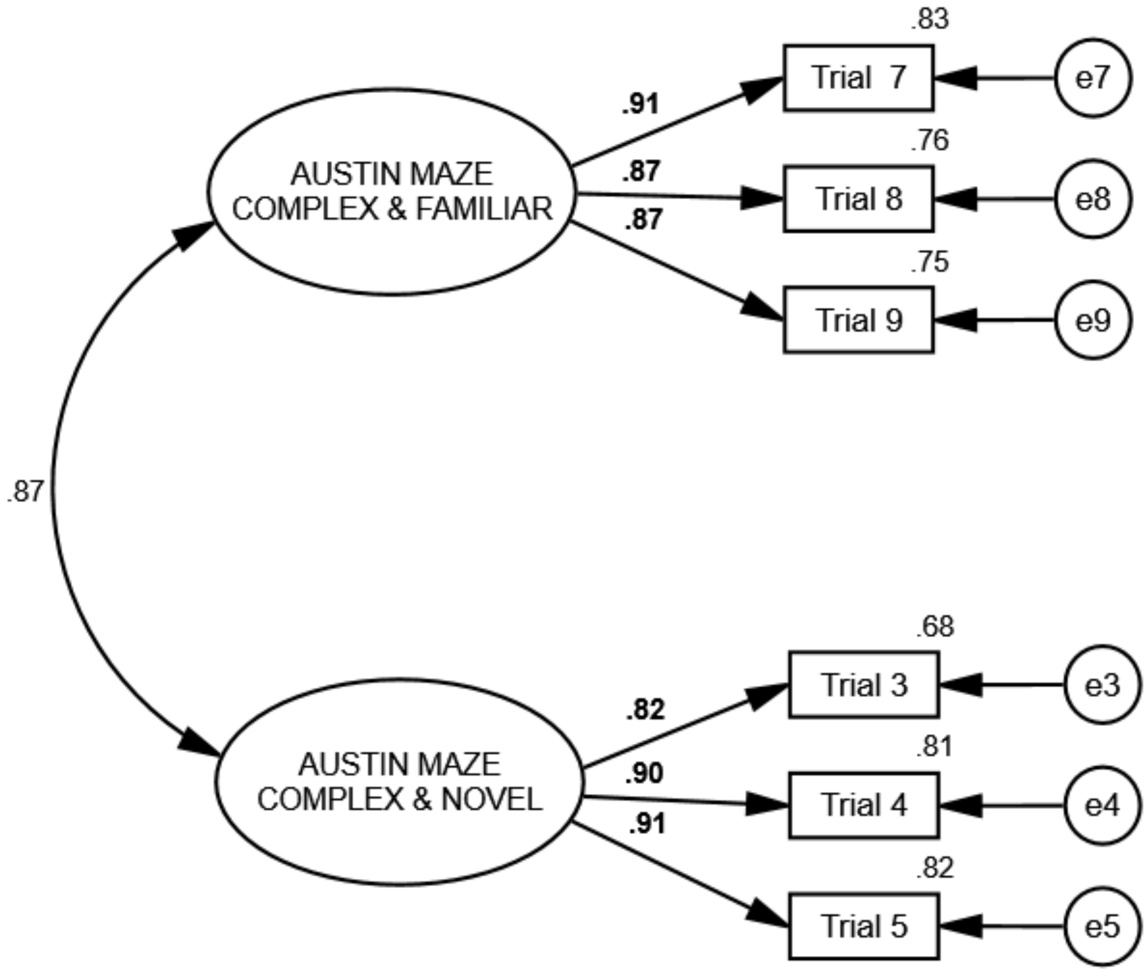
[a] Bootstrapped adjusted standard error.

[b] Unstandardised residual error variance associated with each trial.

A paired sample t-test was conducted to compare mean performance scores between Austin Maze C&F and Austin Maze C&N GDC Factors. There was a significant difference in the mean score for AM C&F (*M*= 24.87, *SD*= .2.89) and AM C&N (*M*= 22.47, *SD*= 3.17) latent factors; $t(102)= 12.594$, $p= <.001$. The final model, including coefficients in their standardized form is illustrated in Figure 7.

Figure 7

Final Model from CFA of Block Design GDC Factor Structure



Note. The number adjacent to the curved double-headed arrow is the inter-factor correlation. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each exogenous variable are the squared multiple correlations (SMC).

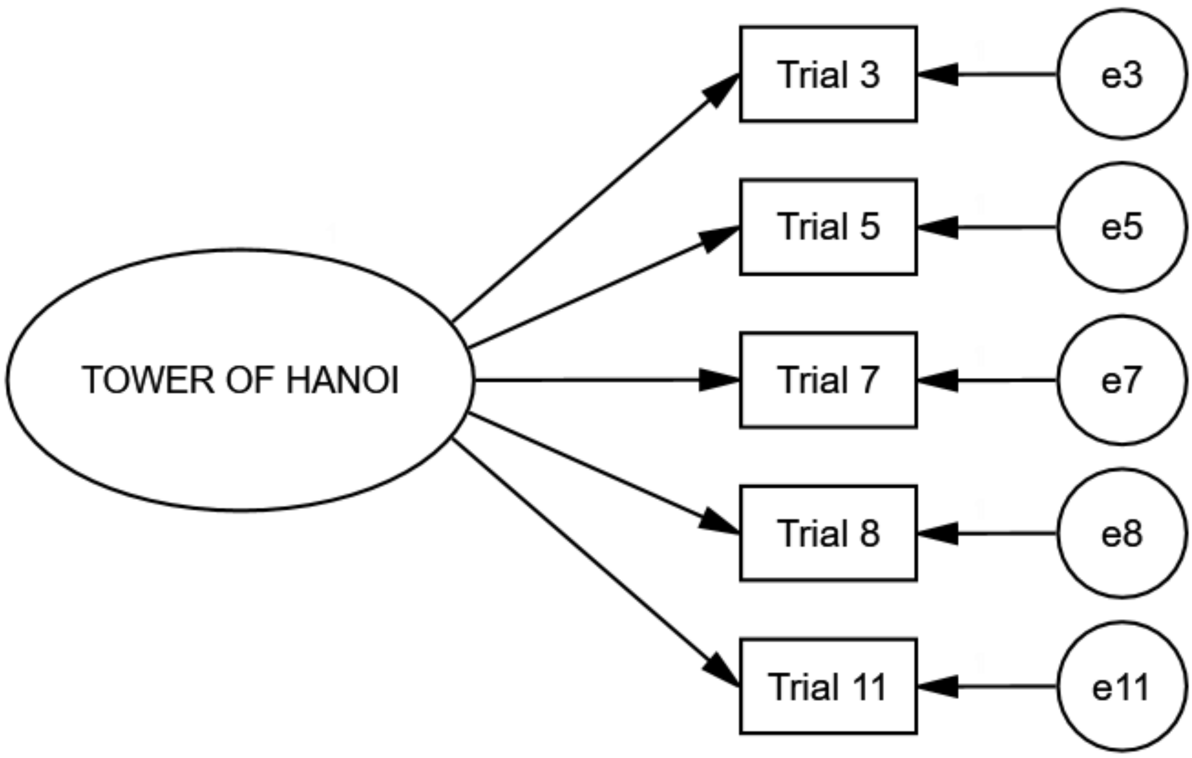
7.2.4 Tower of Hanoi

A CFA was performed using IMB SPSS AMOS v25.0. to investigate whether hypothesised degrees of demand varied within the TOH task. The hypothesised model shown in Figure 8 examined whether one Global Demand Classification from the DCS underpinned performance across the duration of the TOH.

The five-item one-factor structure identified using EFA during Study 1a was then subjected to a CFA using MLE using IMB SPSS AMOS v25.0 to assess whether the hypothesised factor provided a good fit to the sample data.

Figure 8

Hypothesised Model for TOH GDC Factor Structure



Note. Hypothesised model for FAS factor structure. The large circle represents the endogenous latent variable, squares represent exogenous indicator variables, and the small circles represent the residual variance.

7.2.4.1 Assumptions

The assumption of multivariate normality was re-assessed given a potential change in the distribution of performance scores across five trials in comparison to the initial 11-trials. M-distance did not exceed the critical x^2 for $df = 5$ (α=.001) of 20.51 for any cases, indicating that multivariate outliers were not of concern. Mardia's estimate of multivariate normality (7.712) indicated the potential violation of multivariate normality (Table 28). Due to this suspected violation, BS-*p*, and bootstrapped standard errors were requested to adjust for any distribution misspecification of the model.

Table 28

Descriptive Statistics for TOH Variables

				Range			
Variable	*N*	*M*	*SD*	Potential	Actual	Skewness	Kurtosis
TOH 3	103	4.34	4.36	1- ∞	1.00-18.00	1.437	1.113
TOH 5	103	4.74	4.17	1- ∞	1.00-18.00	-0.955	0.034
TOH 7	103	4.20	3.95	1- ∞	1.00-17.00	1.275	0.714
TOH 8	103	3.25	3.30	1- ∞	1.00-13.00	1.316	1.885
TOH 11	103	11.44	13.46	1- ∞	1.00-53.00	1.727	2.296
Mardias Estimate							7.712

Note. ∞ = No maximum score is set by the task. TOH = Tower of Hanoi.

7.2.4.2 Model Estimation

The hypothesised model demonstrated a significant overall good fit to the data x^2(5, *103*) = 4.267, *p*= .512, BS-*p* = .459, SRMR = .416, RMSEA = .000, CFI = 1.000, GFI = .983. An evaluation of the standardised residuals did not reveal any aberrant values that would indicate poor fit (Appendix E). As shown in Table 29, MLE of the model demonstrated significant relationships between all indicator variables and the latent factor. However, residual variances were high, particularly for Trial 11 (Table 29).

Table 29

Maximum Likelihood Estimates for Confirmatory Factor Analysis of TOH

	Estimates				
Parameters	Standardised	Unstandardised	SE[a]	Residual[b]	*p*
Trial 3	.442	1.923	.137	15.207	< .001
Trial 5	.550	2.287	.139	12.028	< .001
Trial 7	.410	1.617	.150	12.907	< .001
Trial 8	.648	2.121	.117	6.289	< .001
Trial 11	.309	4.139	.120	162.494	.012

Note. SE= Standard Error; TOH= Tower of Hanoi

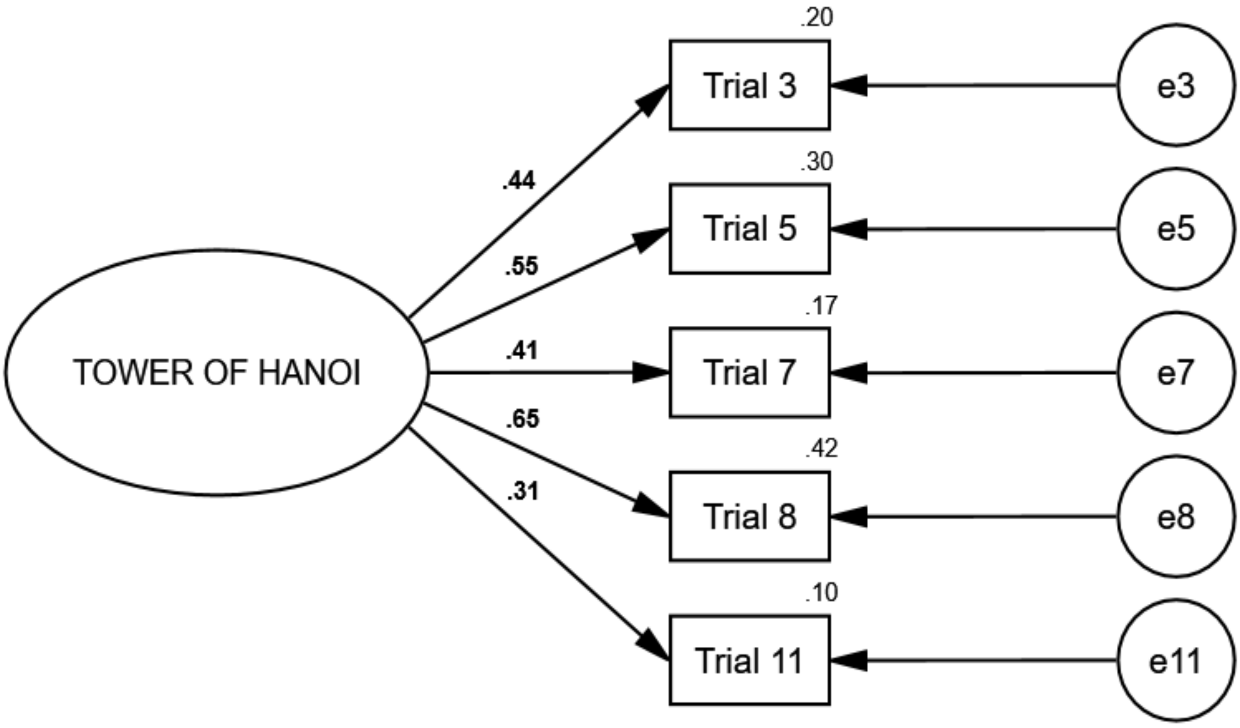
[a] Bootstrapped adjusted standard error.

[b] Unstandardised residual error variance associated with each trial.

SMC values were below the recommended 0.2 threshold for Trials 7 and 11, indicating a poor amount for variance explained in these items by the latent factor (Figure 9). However, the outcome of this CFA was to inform the creation of a weighted composite variable for this latent factor during later studies. Thus, the pooling of the unique explained variance from the indicator variables during this process collectively provided a better measure of the latent factor in comparison to each indicator variable independently. Thus, the items were retained for analysis in *Study 2*. The final model, including coefficients in their standardized form is illustrated in Figure 9.

Figure 9

Final Model from CFA of TOH GDC Factor Structure



Note. The number adjacent to the curved double-headed arrow is the inter-factor correlation. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each exogenous variable are the squared multiple correlations (SMC).

7.2.5 Summary of GDC Measure and DCS Criteria

The results of the current study produced revisions to the DCS Criteria and GDC for task elements that were reported in Study 1a. An update to these revisions to reflect the outcomes of Study 1b is provided in Table 30 below.

Table 30

Summary and Revisions of GDCs and DCS Demand Criteria

Task	1.Abstraction	2.Contextual Stability	3.Action Rules	4.Instructions and rules	5.Duality	6.Automaticity	7.Schematic Demands	8.Episodic Demands
Simple & Familiar Demand Model								
Stroop Test- Word Trial	S	S	S	S	S	F	F	F
TMT- A	S	S	S	S	S	F	F	F
FAS S&F	S	S	C	S	S	F	F	N
Map Search 0-60secs	S	S	S	S	S	F	F	F
5-Point Test 0-60secs	S	S	C	C	S	F	F	F
Test of d2	S	S	S	S	S	N	F	F
Simple & Novel Demand Model								
Visual Span-Backwards	S	S	C	S	S	N	N	N
Digit Span-Backwards	S	S	S	S	S	N	N	N
FAS S&N	S	S	C	S	S	N	F	N
5-Point Test 61-120secs	S	S	C	C	S	N	F	N
Visual Elevator	S	S	S	C	S	N	N	F
Complex & Familiar Demand Model								
Map Search 61-120secs	C	C	C	S	S	F	F	F
TOH	C	C	C	C	S	N	F	F
Austin Maze – C&F	C	S	C	C	S	N	F	F
TMT- B	C	C	C	S	S	N	N	F
Bock Design C&F	C	C	C	S	S	F	F	F
Complex & Novel Demand Model								
Block Design C&N	C	C	C	S	S	N	N	N
Stoop Test- Colour Word	C	C	S	C	S	N	N	F
Austin Maze C&N	C	S	C	C	S	N	F	N
Elevator Counting Reversal	C	S	C	C	C	N	N	F

7.3 Discussion

The aim of Study 1b was to mathematically validate the DCS for tasks traditionally represented by a singular outcome score, even though research has arguably demonstrated that successful performance is underpinned by more than one skill. The tests under investigation were the BD, AM, FAS, and TOH. The outcomes of analyses will first be discussed in relation to each test, followed by a discussion of the overall findings from study 1a and 1b as a collective.

7.3.1 Block Design

It was hypothesised that performance within the BD task would reflect a three-factor model representative of the GDC identified using the DCS during Study 1a. This hypothesis was partially supported, with a three-factor GDC model representing performance variance during S&F, C&F, and C&N conditions of the BD task being found to fit the data. However, in order to ensure parsimony, alternative model comparisons were conducted that revealed a two-factor demand model within the BD task to provide a superior fit. Assessment of the two-factor demand model against the DCS criteria revealed that the later trials shared more conceptual commonality than was first anticipated. The collapsing of the two hypothesised complex GDC factors into one (C&N) highlighted this communality to be shared amongst all nine-block trials of the BD task. This made conceptual sense in consideration of previous research that has linked performance within BD to the increase in uncertainty that a larger quantity of block apparatus can create (Cardillo et al., 2017; Royer et al., 1984; Rozencwajg & Corroyer, 2001). Application of the DCS to the respecified two-factor demand model revealed that the earlier trials (Trials 5, 6, 7, 8, & 9) in the BD test represented C&F demands, and later trials (Trials 10, 11, 12 & 13) were found to reflect C&N demands.

As both factors were considered to reflect variance in performance under (C) Complex Global Demands, the marked difference between the two sets of trials was the high degree of novelty. (N) Novel demands were introduced by changes to the quantity of the blocks and the design configuration that helps to imply an imposed grid. As discussed previously, during the C&N Trials 10 & 11, a new larger square grid of a nine-blocks is presented. Trial 12 then continues to introduce novel demands by rotation of the design grid into a diamond configuration. Lastly, during Trial 13, a novel problem is introduced whereby the border of the design grid is no longer defined, leaving the individual to determine the correct configuration (square or diamond). In addition to the changing grid configuration, later trials of the BD require the increasing uptake and use of the split-half (red & white) sides of the blocks, with Trial 13 utilising these completely. This is in contrast to C&F Conditions where solid coloured sides are used more frequently, under a square imposed grid. The additional uptake and use of split coloured sides has previously been linked to task uncertainty amongst BD paradigms (Royer et al., 1984), which can create a negative impact on BD performance. The impact of this introduction of (N) Novel demands during later trials within the BD test was reflected by the significant decrease in performance scores that was found in comparison to the C&F trials.

The high correlation between C&F and C&N GDC Factors suggests that while the demands within the BD task are divisible both statistically and conceptually, they must be considered and interpreted as highly interrelated constructs. In part, this relationship is likely reflective of the (C) Complex demands that the two factors were found to share. These include shared demands for (1.C) Complex levels of abstraction required to substantiate sub-goals and to evaluate the most effective use of the apparatus to ensure correct block positioning, (2.C) similarity in the Complex contextual stability, and (3.C) comparable Complex demands in the available action rules.

In addition, this relationship may be attributed to task-specific schemas developed during C&F demand trials that support performance during C&N demand trials. Due to the embedded nature of performance of each factor within the same global test, absolute Novelty within the C&N trials cannot be assumed. Instead, it is likely that during the C&F trials of the

BD, fundamental schemas about the block apparatus are established. Performance during the C&N trials may then be supported by these familiar set of task-specific rule representations. During the C&N trials, these previous task-specific schemas may then allow cognitive control to be proactively applied, if familiar contexts, cues and actions rules are sufficiently detected by the individual during the C&N trials (Verbruggen et al., 2014). Thus, performance during C&N Demands of the BD may represent several challenges to cognitive control, whereby performance is influenced by (a) how well the fundamental BD schemas are established, (b) the ability to detect similarities in environmental cues to enable preparation for the anticipated S→R, and (c) the ability to manipulate and update schemas to novel task-specific requirements.

7.3.2 Austin Maze

It was hypothesised that performance within the AM would reflect a two-factor model representative of the GDCs identified via the DCS during Study 1a. This hypothesis was supported, with a two-factor demand model accepted to provide the most superior fit over alternative models. The final model supported the GDC for the AM and represented the C&F demands of the three later trials (Trials 7, 8 & 9), and the C&N demands of the three earlier trials (Trial 3, 4, & 5).

A two-factor solution to the performance variance in the AM demonstrates that task demands can vary, even in the absence of any task exogenous administrational change, and that these demands can indeed influence task performance. Alternative models that were analysed included Trial 2 and Trial 6 as an indicator variable, which were rejected due to resulting in a poorer model fit. Trial 2 was considered to reflect the initial exposure and discovery of the hidden path, whereas Trial 6 was thought to reflect a period within the task where novel demands begin to subside and transition to familiar demands. Thus, the improvement of model fit with the exclusion of both trials likely reflects the increased strength in communality that can be explained by each GDC factor.

High correlations were found between the C&F and C&N GDC Factors. Similar to the findings within the BD task, the embedded nature of the trials and mutual similarity in their (C) Complex demands likely represents this shared relationship. Across the AM task complexity can therefore be considered somewhat stable due to (1.C) high levels of abstraction required between the hidden path, the feedback that is provided by the apparatus, and the execution of a response; (2.S) similarity in their contextual information via the application of a stable maze configuration; (3.C) comparable Complex demands in action rule complexity that require the participant to select the correct tile from numerous available alternatives; (4.C), the maintenance of a Complex set of multiple rules and instructions that are task centred; and (5.S) the absence of any dual-task demands across both sets of trials.

The divergence of the GDC factors can predominantly be attributed to the change in the (N) Novelty of demands across the task administration. The C&N trials (Trials 3, 4, & 5) of the AM encompass initial novel learning of the task that require (4.C) multiple sets of instructions to be followed; (3.C) while determining the correct action rules; (6.N) in a novel controlled setting; (8.N) all while aiming to maintain the new path in memory. During later (C&F) trials of the AM, familiarity of the hidden path increases from repeated exposure, and (7.F) schemas for the maze are required to be formed to enable completion of the correct path. Moreover, (8.F) Episodic requirements should be reduced due to repeated exposure of the correct path over across multiple trials. The impact of these changing demands within the AM was further highlighted by a significant decrease in the number of incorrect tiles that were selected during the later C&F trials. Given the purpose of the AM is to discover the hidden path, gaining familiarity with the task is a primary objective. Thus, a two-factor GDC structure of the AM allows for an understanding of where (N) Novel and (F) Familiar demands become distinguished. Failure to retain episodic information from the task, and form schematic relationships could result in (N) Novel demands remaining high, which may result in a higher number of incorrect tiles being selected during later trials. Conversely, the early retention of episodic information and formation of schematic relationships could result in a decrease in incorrect tiles being selected during earlier trials.

The two-factor GDC model of the AM offers an empirical insight into the variety of demands that exist within the task, which aligns with previous research that has proposed various cognitive skillsets to be recruited during hidden maze test performance (Crowe et al., 1999). The application of the DCS and the subsequent support for its factor structure provides a useful framework to represent changes in trial performance that reflect different demands for cognitive control and the recruitment paradigm specific cognitive skillsets.

7.3.3 FAS Test

It was hypothesised that performance within the FAS test would reflect the two GDCs identified via the DCS for the FAS during Study 1a. The hypothesis was supported, with a two-factor model accepted to represent performance during S&F (F' 0-15secs, 'S' 0-15secs, 'F' 15-60secs), and S&N demand conditions ('F' 16- 60secs, 'A' 16- 60secs, 'S' 16-60secs).

Similar to the AM paradigm, the identification of distinguishable performance variance supports the application of the DCS to the task, in addition to the need to recognise that task demands within the FAS test change in the absence of any exogenous alterations to task administration. The high correlation between the two GDC Factors of the FAS likely reflects a predominant communality in the (S) Simple demands that the two factors share. These can be attributed to (1.S) low levels of abstraction required due to the direct interaction with the task instructions and the response required; (2.S) the stable presence of contextual information and explicit notification that is given to prompt the letter change; (3.C) comparable demands in action rule complexity that require the participant to produce the correct word from those available in their semantic memory; (4.S) consistent and unchanging instruction and rules across the task; and (5.S) absence of any dual-task demands across both sets of trials.

As both demand factors were considered to reflect the (S) Simple demands within FAS test, the marked difference between the two sets of trials was the increase of (N) Novel demands during later stages ('F' 16- 60secs, 'A' 16- 60secs, 'S' 16-60secs). The (N) Novel demands of the FAS reflect the need to (6.N) exert greater control over word production due to the exhaustion of words previously available in semantic memory; and (8.N) the increase in episodic demands due to the need retain an increasing quantity of previous words to ensure repetition is avoided. Conversely, during the initial 15-second quartiles of the task (S&F), initial word generation is likely to be more simply (6.F) dependent on the immediate recall of familiar known words stored in semantic memory (Venegas & Mansur, 2011). This relationship may also represent how performance during S&F trials can influence performance during later (N) Novel conditions. For example, familiarity within the FAS test may potentially extend into the novel conditions when an increase in the number of schemas for words, and overall vocabulary is held by the individual. Therefore, the effects of (N) Novel demands may have the potential to be mitigated. Alternatively, given the embedded nature of the S&F demands within the FAS test, failure to execute S&F Demands (e.g. the lack of words in semantic memory) may have a subsequent impact on performance during (N) Novel Demand conditions.

The identification of (S) Simple Demands within the FAS was interesting given that the FAS test is commonly used as a measure of the EF, which by common definition is elicited predominately during complex demands. According to the current findings, it is likely that the demands within the FAS elicit controlled behaviour due to the presence of (N) Novel demands during later stages of completion only. This finding adds support for previous suggestions that during later stages of the FAS the EFs of planning and monitoring of performance are required (Venegas & Mansur, 2011).

7.3.4 Tower of Hanoi (TOH)

The one-factor model (Trials 3, 5, 7, 8, & 11) of the TOH demonstrated a good fit to the performance data. However, despite being statistically significant, low variance was explained for Trial 11. This trial differed qualitatively from the other trials within the factor by the increase of one additional disk apparatus, and the increase in the number of sub-goals that were required to be established for successful performance. These additional requirements may represent emerging demands for novelty within the TOH for this trial. This was further highlighted by the high residual variance of Trial 11. However, the variance that was explained in Trial 11 was significant in the presence of the overall model. Thus it was retained and considered suited to the overall fit of the TOH C&F demand model.

Assessment of the TOH using the TowerTool v2.0 enabled insight into the TOH parameters that were shared amongst the trials that the latent factor represented. Interestingly, all trials within the factor comprised of a different number of possible minimum moves, thus contrasting previous research that attributed complexity of tower trials to this notion (Berg et al., 2010; Zook et al., 2004). All five trials were found to have the same starting Tower configuration, however this was the only consistent commonality shared that was identifiable using the TowerTool v2.0. A majority of trials (4 out of 5) were flat ending, with the exception of Trial 8. Another prominent commonality was with the similar number of subgoals required by each trial. While the purpose of this study was not to uncover the most influential contribution of tower configurations, the current findings provide an insight into what parameter communalities can be shared between trials, and that these communalities can hold both statistical and conceptual significance. This finding serves to highlight that similarities can be shared amongst Tower Trials, however they appear to reflect a multifaceted contribution to the overall complexity of the trial as a GDC Factor.

Application of the DCS revealed that the singular TOH factor represented C&F demands. Given that the latent factor represented a spread of early to late TOH trials, the factor was considered to represent the performance capacity to maintain and execute the task under TOH specific (F) Familiar demands. The (F) Familiar demands of the TOH are likely reflected by the need to approach a new Tower starting configuration while (7.F) maintaining the global schemas formed during the TOH trials (E.g. move the larger disc to its goal state first). This is also extenuated by the similarity in starting and goal positions that are shared between the C&F trials. The (C) Complexity of the TOH GDC Factor reflects the need for (1TPR.C) complex abstraction between the start and goal state, and formulating the appropriate sub-goals, evaluation of actions between start and goal states, and the relationships between the apparatus, goal state and planned movements; (2.C) changes in contextual demands between each trial, with both configurations and disc numbers changing; (3.C) the complex set of action rules available by which the best solution must be determined by the participant; (4.C) the need to maintain multiple task-specific rules across all trials during each (5.S) single trial completion.

Overall, the results support that the TOH is heterogeneous and multifaceted in its demands, but communality can exist between five-TOH trials due to similarities in their inner problem structure and demands. Acceptance of the TOH C&F GDC Factor serves as a useful testing parameter for research or clinical settings to better understand the impact of these demands in the context of the TOH paradigm. The failure of the remaining TOH trials to converge into any additional significant factor structure further highlights the potential heterogeneity of demands that the remaining TOH trials may encompass. It would be of interest for future research to explore GDC Factor structure amongst a Tower task with additional trials that may offer an increased variety of parameters and testing demands, compared to the version that was used here.

7.3.5 General Discussion of Study 1

A valuable methodological approach of the current study was to guide test selection, task element analysis, and demand classifications by previous literature, neuroscientific evidence, and demand features, but with less regard for established purported terminology or definitions. Each task was defined and assessed at the micro-level using the DCS in Study 1a, and then further supported mathematically by Factor Analysis here in Study 1b. Essentially, this study attempted to take the advantages of the cognitive control theory approach to task analysis, and apply it to the more generalised EF assessment approach that remains wedded to the notion that task performance is equal or exclusive to a component of EF.

Previous research has been successful at identifying which predominant EFs may be shared amongst a set of assessment tasks. However, this approach fails to offer context to performance in relation to the controlled behaviours that are required to successfully respond to the demands that are established by the task. This study has demonstrated that Barkley's (2012) contention that current neuropsychological measures only allow for an understanding of the components of EF, but not its adaptive nature, cannot be doubted.

Traditional outcome scoring derived for BD indicate visuo-constructional accuracy as a function of speeded performance, averaged across numerous trials of *varying difficulty.* The finding that BD is represented by two separable GDC (C&F, and C&N) in the current study does not replace that score and *what* it represents. However, when used with the DCS, scores will instead reflect *what* and *why* people perform differently relative to task demands.

In contrast, the finding that TOH yielded only one GDC across specific trials enforces the notion that the TOH, considered as one overarching measure of a single construct (e.g. planning), is confounded by significant "noise". This provides strong evidence for the task impurity inherent in this task, which authors have argued for decades perpetuates the failure of EF theory to maintain ecological validity. The TOH findings also enforce the current contention that GDC classification should not be used as a replacement for EF test performance, but should be used as an adjunct. For example, the TOH C&F classification does not mean that the trials that comprise it should be isolated and administered in the absence of the other trials. In fact, the identification of this ‘purer’ representation of shared characteristics relies on the administration of earlier TOH trials. What is indicated strongly, however, is that averaging performance across all trials and providing a single ‘planning score’, as per traditional measurement paradigms, will include a range of non-planning skills.

To compare the outcomes of BD and TOH, the former has identified essentially varying complexity of demand across the entire task, whereas the TOH has revealed an isolated shared variance with a degree of ‘noise’. Thus, the measures themselves must be divided into conditions that reflect their level of complexity and novelty, making them more consistent with both neuroscientific EF imaging findings and cognitive control theory if the clinical utility of EF tests will endure.

It could be argued that the mathematical analysis performed during Study 1b is somewhat self-fulfilling of the outcomes of Study 1a. However, in contrast to previous factor analytical research, with the exception of the TOH, the structure of the models analysed during Study 1b were theoretically informed, instead of being entirely data driven. This approach was only possible due to the development and application of the DCS framework in Study 1a, which was applied without reference to trial or task scores. The statistical support demonstrated during Study 1b therefore offers preliminary support for the acceptance of the DCS as a suitable tool for identification of demands for complexity and novelty within multifaceted tests of cognition.

Notably, it was demonstrated across all models that performance was better represented when the GDC were taken into consideration, instead of tasks being treated as singular in their outcomes. This finding supports previous neurological and behavioural studies that have correlated task performance changes with changes to the demands for cognitive control within multifaceted testing environments (Baker et al., 1996; Connolly et al., 2016; Fedorenko et al., 2013; Jaeggi et al., 2003; Konishi et al., 1999; Lie et al., 2006; Unterrainer et al., 2004; Yoshida et al., 2010).

While these studies were successful in highlighting performance and neurological activation differences at various stages during test performance, the collective findings of Study 1 offer a further capacity to identify both *where* and *why* performance may vary in response to demand. The provision of the DCS enabled the identification of *where* key changes in both complexity and novelty can present within a test environment that would, in turn, call for changes to the recruitment of cognitive control resources. Concurrently, the appraisal via the DCS against the demand features for cognitive control provided an understanding towards *why* demands may change. Study 1b demonstrated the success of this approach to achieve greater clarity of performance variance that occurs within multifaceted tests. The provision of this approach enabled performance to be interpreted in relation to the changes to demands for cognitive control, instead of the attributing performance to a unitary cognitive skillset.

The next step in the evaluation of complexity and novelty, and their interaction, is the analysis of between task convergence. This approach mirrors previous efforts (Anderson, 2002; Messer et al., 2018; Miyake et al., 2000; Testa et al., 2012) where EF tests are investigated to load onto one or more factors, but rather than traditional purported EFs and their scoring, this study has facilitated an investigation of whether tasks with similar demands load together, irrespective of the skill that underpins them.

Chapter 8

Study 2: Exploring the GDC

Across the four multifaceted tasks analysed during Study 1b, all four DCS Global Demands were identified. With the exception of the first 15-seconds of the FAS which was found to encompass S&F demands, GDC Factors were classified to remain relatively stable across the continuum of complexity, with any marked differences being attributed to the continuum of novelty that exists within their test environments. The stability in complexity is likely attributed to the fact that the basic parameters of a task don't change when novelty is applied to increase difficulty. Therefore, the findings suggest that performance during (F) Familiar demands of the task may influence, or even predict, performance during (N) Novel demands. The importance of this relationship was demonstrated by the high correlations found within test administrations that featured both (F) Familiar and (N) Novel Global Demands. The finding that a continuum of complexity can remain stable within a task, but novelty can vary, demonstrates support for the main hypothesis of this project that both C&N demands exist on separable, but intersecting dual axes. The further exploration of this relationship is warranted in order to better understand the nature of the relationships between (F) Familiar and (N) Novel Global Demands. Moreover, the relationship between (S) Simple and (C) Complex Global Demands also requires further exploration as neither were found to feature together within the same test.

The aim of Study 2 was to explore the communality that is shared by each GDC between different neuropsychological tests. This will help to determine whether task elements classified as similar in complexity and novelty hold together mathematically, which will allow comparisons not between tests but across demands. GDCs that were identified for neuropsychological tests during Study 1a and 1b were used to establish four hypothesised GDC models; S&F, S&N, C&F, C&N.

8.1 Method

Weighted composite scores were computed for each participant to provide an estimated performance score for each GDC that was identified during Study 1b. This approach enabled direct reference to the GDC performance scores of that task instead of retaining each indicator variable from all analyses. Adopting this weighted approach enabled a proportional composite score to be established that better represented the unique shared variance that was explained by each GDC factor (Jöreskog & Sörbom, 1989). Factor score weights were obtained using IBM SPSS AMOS v25.0 for each GDC Factor from Study 1b (Appendix K).

The formula used to calculate weighted composites for each latent factor is as follows:

$$\xi = \omega X$$

Note:

ξ = composite score
ω = factor score weights
X = participants observed indicator variable scores

As factor score weights are proportional weights, all factor weights were rescaled to total one (1). Re-scaling the composites in this manner allowed for the composite to exist on the same scale as the indicator variables. To achieve this, each factor score weight was divided by the total sum of factor score weights for each GDC factor. These newly calculated composite scores were given a revised label that included the GDC that they represent. For example, the new label ‘BD C&F’ represented GDC performance variance encompassing Trials 5-9 of BD.

8.1.1 Congeneric Model-Based Measures of Reliability

All accepted GDC models were subject to a reliability analysis. Traditional approaches to assessing reliability often do not take into account measures being congeneric (Terry &

Kelley, 2011). To provide an accurate assessment of reliability of each congeneric model computed with the current study, Hancock & Mueller's (2001) Coefficient *H* was calculated. The advantage of Coefficient *H* formula was its ability to maximise composite reliability by taking into account the contribution of all variables. This is in contrast to other measures of reliability, such as Cronbach's alpha, which is often an underestimation of congeneric measure due to the assumption of equal parameter contributions to a model (Widhiarso & Ravand, 2014). Coefficient *H* ranges from 0-1, with 1 representing complete collinearity between variables, and 0 representing zero factor loadings. Hancock & Mueller (2001) suggest a good indication of reliability is when a coefficient of .70 or greater is reported for one or two standardised loadings, or a coefficient of .60 and above for three or more standardised loadings.

The formula for coefficient *h* was as follows:

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

Note: λ's are the standardised factor loadings.

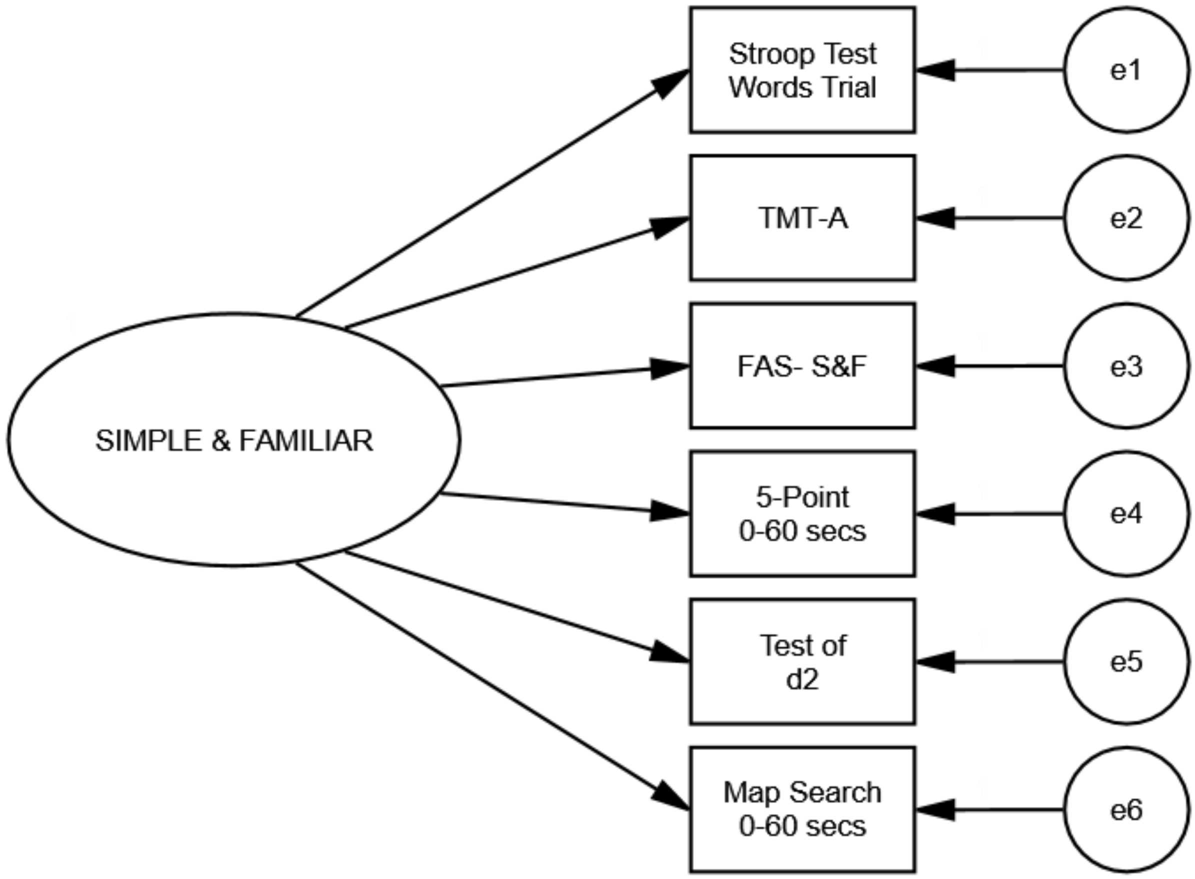
8.2 Results

8.2.1 Simple and Familiar

A CFA was performed using IBM SPSS AMOS v25.0 to investigate whether neuropsychological test and task elements identified via the DCS to comprise of a S&F GDC represented a 1-factor model solution (Figure 10).

Figure 10

Hypothesised S&F GDC Model



Note. FAS S&F = FAS Simple and Familiar Score; TMT= Trail Making Test.

The larger circle represents the endogenous latent variable, squares represent exogenous indicator variables, and the small circles represent the residual variance.

8.2.1.1 Assumptions

As shown in Table 31, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality, and Mardia's estimate was within the acceptable range to support the assumption of multivariate normality. M-distance did not exceed the critical x^2 for df = 5 (α= .001) of 20.51 for any cases, indicating that multivariate outliers were not of concern. There were no missing data.

Table 31

Descriptive Statistics for S&F GDC Model Variables

Variable	*N*	*M*	*SD*	Range: Potential	Range: Actual	Skewness	Kurtosis
Stroop Test – Words[a]	103	7.79	2.42	0 - ∞	0.250 - 12.10[a]	-0.689	0.156
TMT A[a]	103	19.80	6.58	0 - ∞	0.430 - 31.44[a]	-0.762	0.005
FAS S&F	103	4.17	1.08	0 - ∞	1.58 - 7.49	0.293	0.399
5-point 0-60secs	103	16.30	4.77	0 - 80	5.00 - 26.00	-0.201	-0.546
Test of d2	103	179.42	37.14	0 - 299	80.00 -257.00	-0.423	0.081
Map Search 0-60secs	103	40.61	11.56	0 - 77	19.00 - 69.00	0.293	-0.642
Mardia's estimate							-.419
M-distance							15.52

Note. TMT A = Trail Making Test - A; FAS S&F= FAS Simple and Familiar score. ∞ = No maximum score is set by the task.

[a]Values displayed are an inverse of the original score due to rescaling procedures (Section 5.2.15).

8.2.1.2 Model Estimation and Specification

A CFA using MLE was performed using data from 103 participants. The independence model that tests the hypothesis that all variables are uncorrelated was rejected, $x^2(10, 103) = 61.847$, $p = <.001$. The hypothesised model was then tested and demonstrated a good fit to the observed data $x^2(9, 103) = 4.563$, $p = .871$, SRMR = .0381, RMSEA = <.001, CFI = 1.000, GFI = .985, AIC = 28.563. However, the coefficient predicting the *Map Search 0-60secs* score was non-significant ($p = .073$). Consequently, this variable was excluded and the model re-estimated, $x^2(5, 103) = 1.107$, $p = .953$, SRMR = .0204, RMSEA = <.001, CFI = 1.000, GFI= .996, AIC = 21.107. The AIC index score indicated a better fitting and more parsimonious model after the

Map Search 0-60secs variable was removed. An evaluation of the standardised residuals did not reveal any other aberrant values that would indicate poor fit (Appendix L).

As shown in in Table 32, MLE of the hypothesised model demonstrated significant relationships between all indicator variables and the SIMPLE & FAMILIAR latent factor.

Table 32

Maximum Likelihood Estimates of S&F GDC Model

	Estimates				
Parameters	Standardised	Unstandardised	SE	Residual[a]	*p*
Stroop Test – Words ← S&F	.763	1.834	0.286	2.415	< .001
TMT A ← S&F	.443	2.906	0.740	34.520	< .001
FAS S&F ← S&F	.414	0.448	0.122	0.967	< .001
5-point 0-60secs ← S&F	.397	1.881	0.539	18.857	< .001
Test of d2 ← S&F	.578	21.247	4.265	901.710	< .001

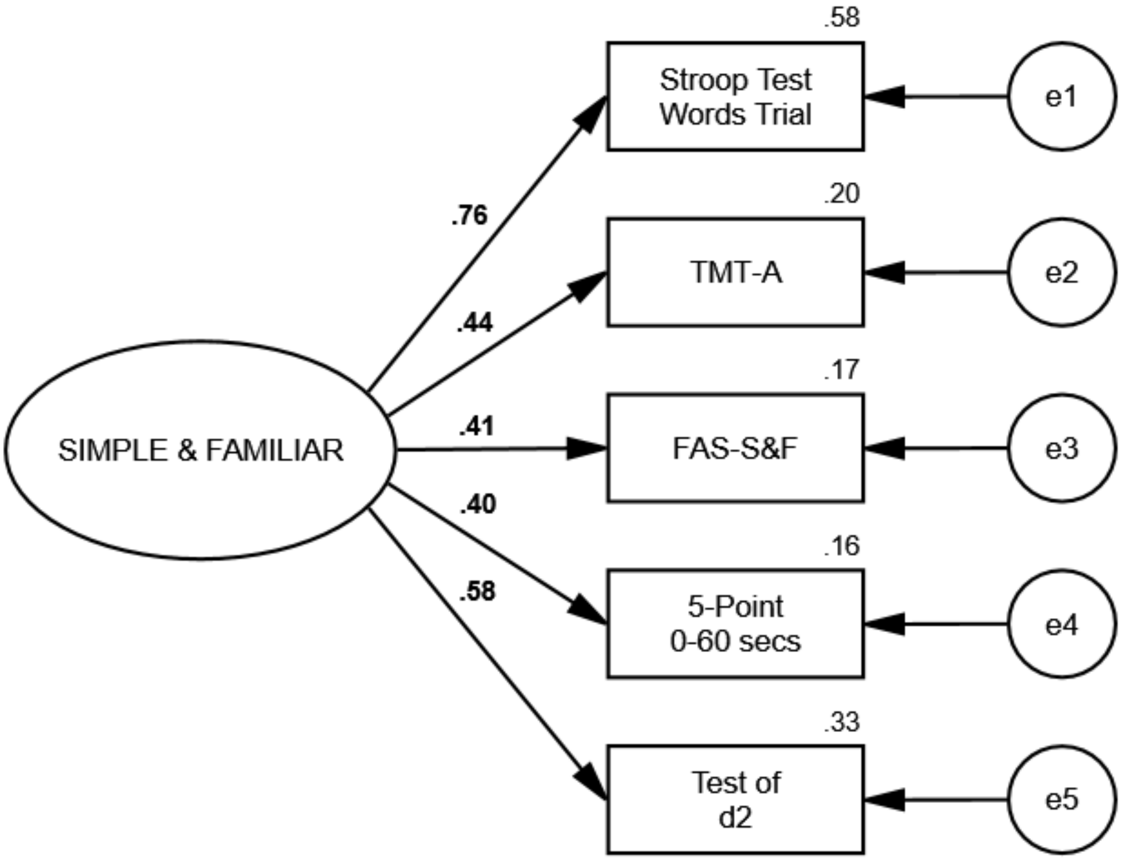
Note. SE= Standard Error; S&F = Simple and Familiar

[a] Unstandardised residual error variance associated with each trial

SMC values (Figure 11) indicated that individual item reliabilities were low overall, with the exception of the *Stroop Test -Words* indicator variable. However, all variables were retained as the aim of this analysis was the identification of a measurement model to determine appropriate factor weightings for the creation of a SIMPLE & FAMILIAR composite score. Coefficient *H* was calculated to test the overall reliability of the model, which indicated an overall good reliability of the SIMPLE & FAMILIAR model (0.71). The final model, including coefficients in their standardised form is illustrated in Figure 11.

Figure 11

Final S&F GDC Model



Note. FAS S&F = FAS Simple and Familiar Score; TMT= Trail Making Test.

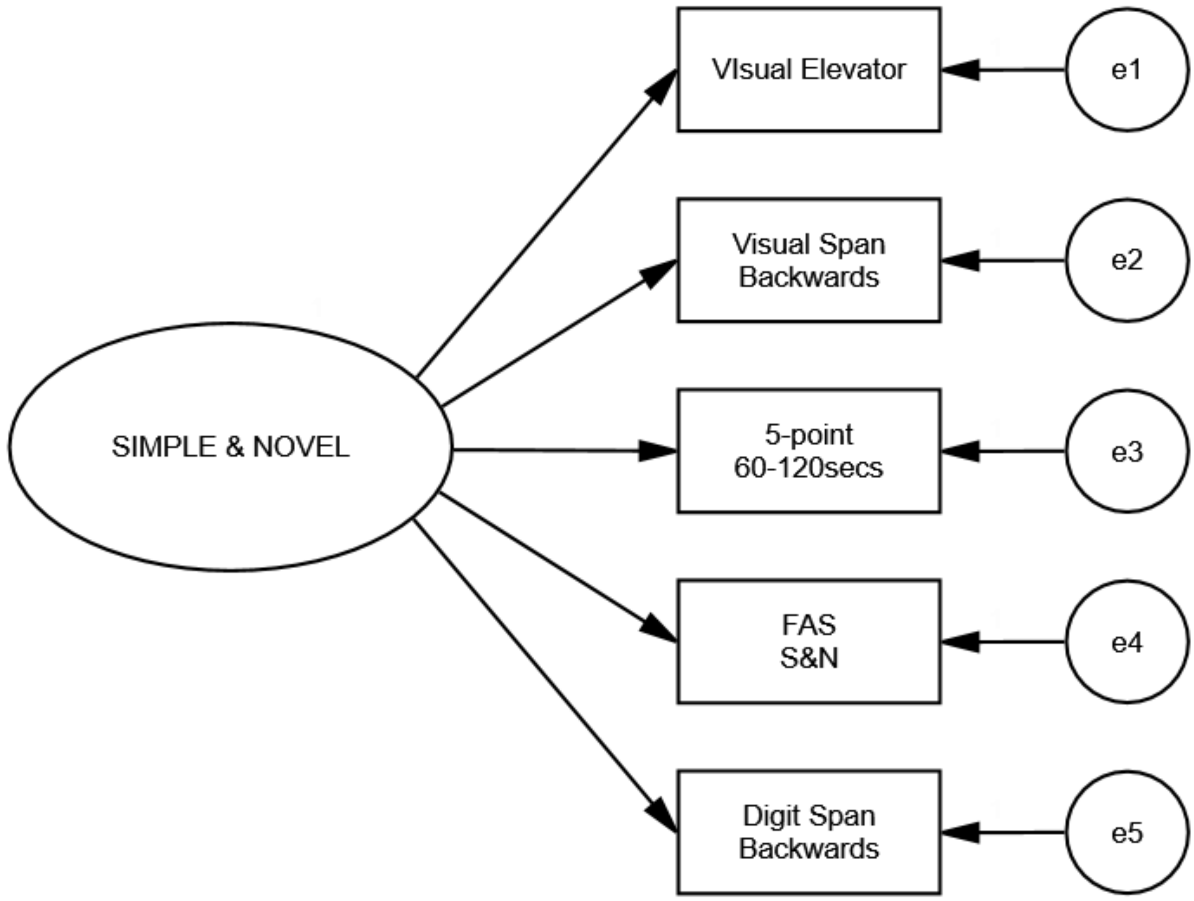
The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each indicator variable are the squared multiple correlations.

8.2.2 Simple and Novel

A CFA was performed using IBM SPSS AMOS v25.0 to investigate whether neuropsychological tests and task elements identified via the DCS to comprise the S&N GDCs represented a 1-factor model solution (Figure 12).

Figure 12

Hypothesised S&N GDC Model



Note. FAS S&N = FAS Simple and Novel Score. The larger circle represents the endogenous latent variable, squares represent exogenous indicator variables, and the small circles represent the residual variance.

8.2.2.1 Assumptions

As shown in Table 33, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality, and Mardia's estimate was within the acceptable range to support the assumption of multivariate normality. M-distance did not exceed the critical x^2 for $df = 5$ (α=.001) of 20.51 for any cases, indicating that multivariate outliers were not of concern. There were no missing data.

Table 33

Descriptive Statistics for S&N Model GDC Variables

				Range			
Variable	*N*	*M*	*SD*	Potential	Actual	Skewness	Kurtosis
Visual Elevator	103	3.84	0.88	0 -13.4	0.70– 5.70	-1.144	1.844
Visual Span-Backwards	103	6.44	1.75	0 - ∞	3.00 – 12.00	0.290	-0.077
5-point 60-120secs	103	9.74	3.11	0 - 80	4.00- 18.00	0.284	-0.665
FAS S&N	103	3.12	0.90	0 - ∞	1.05- 5.50	0.262	-0.298
Digit Span-Backwards	103	5.78	2.39	0 - ∞	1 – 14.00	0.807	0.965
Mardia's estimate							2.289
M-distance							20.123

Note. FAS S&N = FAS Simple and Novel Score.

8.2.2.2 Model Estimation and Specification

A CFA using MLE was performed using data from 103 participants. The independence model that tests the hypothesis that all variables are uncorrelated was rejected, x^2(10, *103*) = 36.972, p= .<.001. The hypothesised model was then tested and demonstrated a good fit to the observed data, x^2(5, *103*) = 3.710, *p*= .592, SRMR = .0404, RMSEA = <.001, CFI = 1.00, GFI = .986. An evaluation of the standardised residuals did not reveal any aberrant values that would indicate poor fit (Appendix M).

As shown in in Table 34, MLE of the hypothesised one-factor model demonstrated significant relationships between all indicator variables and the S&N latent factor.

Table 34

Maximum Likelihood Estimates of S&N GDC Model

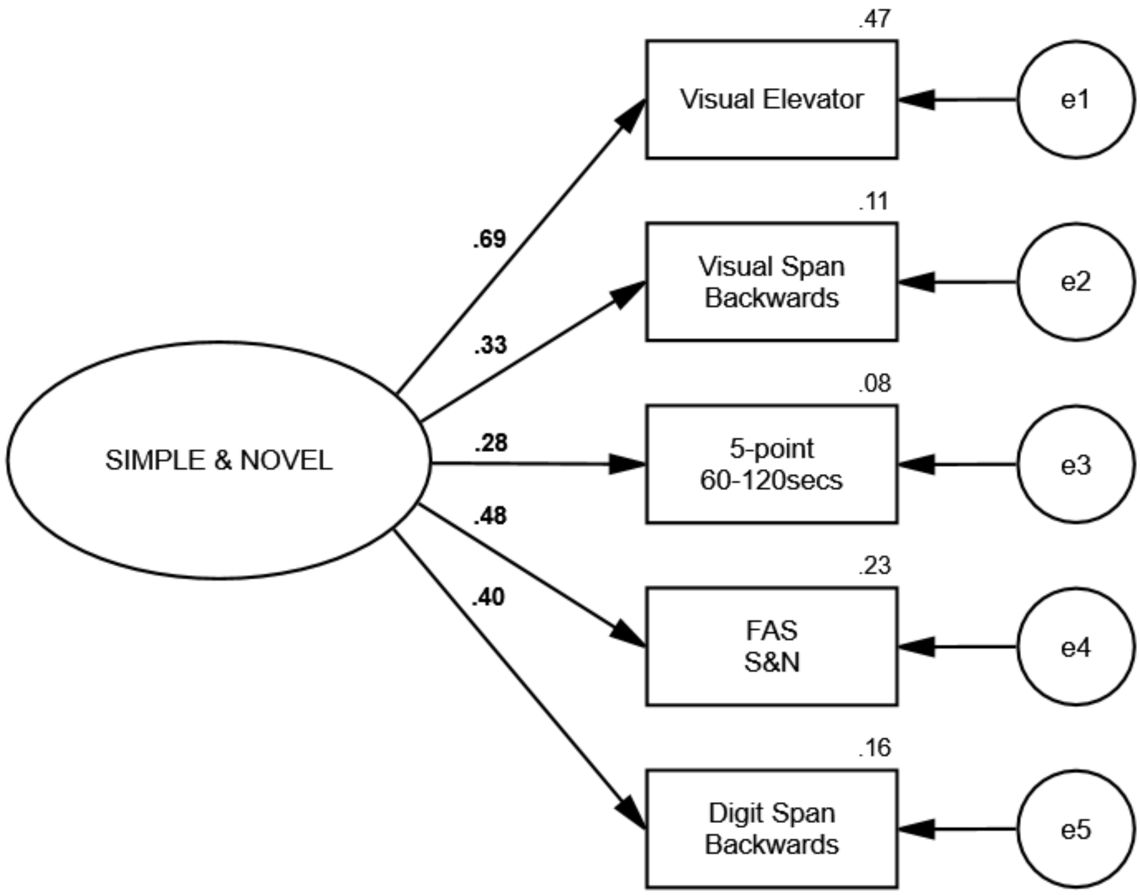
Parameters	Estimates		SE	Residual[a]	*p*
	Standardised	Unstandardised			
Visual Elevator ←S&N	.689	0.604	.129	0.404	<.001
Visual Span Backwards ←S&N	.332	0.580	.220	2.726	.008
5-point 60-120secs ←S&N	.283	0.879	.390	8.852	.024
FAS S&N ←S&N	.482	0.434	.117	0.625	<.001
Digit Span - Backwards ←S&N	.404	0.962	.302	4.737	.001

Note. SE= Standard Error; S&N = Simple and Novel.
[a] Unstandardised residual error variance associated with each trial.

As seen in Figure 13, SMC values indicated that individual item reliabilities were low overall, with the exception of the *Visual Elevator* variable. However, all variables were retained as the aim of this analysis was focused on the identification a measurement model to determine correct factor weightings for the creation of a SIMPLE & NOVEL composite score. Coefficient *H* was calculated to test the overall reliability of the model, which indicated an overall good reliability of the SIMPLE & NOVEL model (0.617). The final model, including coefficients in their standardized form, is illustrated in Figure 13.

Figure 13

Final S&N GDC Model



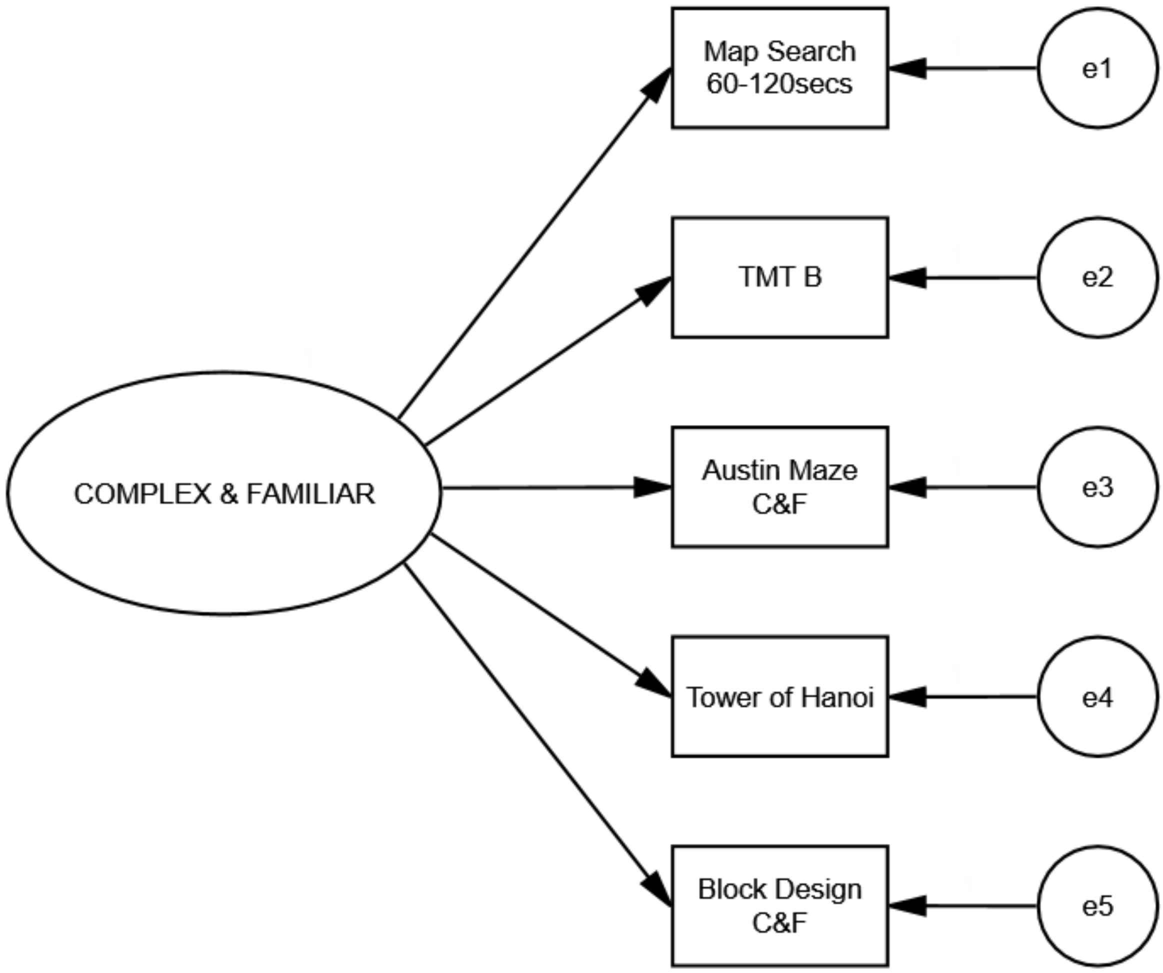
Note. FAS S&N = FAS Simple and Novel Score. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each indicator variable are the squared multiple correlations.

8.2.3 Complex and Familiar

A CFA was performed using IBM SPSS AMOS v25.0 to investigate whether neuropsychological tests and task elements identified via the DCS to comprise the C&F GDC represented a 1-factor model solution (Figure 14).

Figure 14

Hypothesised C&F GDC Model



Note. Austin Maze C&F = Austin Maze C&F Score; Block Design C&F = Block Design C&F Score. The larger circle represents the endogenous latent variable, squares represent exogenous indicator variables, and the small circles represent the residual variance.

8.2.3.1 Assumptions

As shown in Table 35, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality, and Mardia's estimate was within the acceptable range to support the assumption of multivariate normality. However, analyses of M-distance revealed one case that exceeded (M-Distance = 22.10) the critical x^2 for $df = 5$ ($\alpha = .001$) of 20.5. In light of the current sample size adequately meeting conventional requirements for CFA and SEM analyses, the case was excluded from this analysis. Re-computation of M- distances on $n = 102$ revealed no cases that exceeded the critical x^2 for $df = 5$ ($\alpha = .001$), indicating that multivariate outliers were no longer of concern. There were no missing data in the remaining cases.

Table 35

Descriptive Statistics for C&F GDC Model

Variable	*N*	*M*	*SD*	Range: Potential	Range: Actual	Skewness	Kurtosis
Map Search 60-120secs	102	69.00	8.42	0 - 80	46-80	-1.054	0.275
TMT-B	102	61.12	15.80	0 - ∞	12.97-87.56	-0.863	0.971
AM C&F	102	2.03	2.96	Ø	18.95-31.65	-1.397	1.655
TOH	102	7.79	2.67	Ø	0.544- 11	-0.723	0.588
BD C&F	102	5.64	1.59	Ø	2.18 – 8.60	0.050	-0.936
Mardia's estimate							0.777
Maximum M-distance							13.40

Note. TMT= Trail Making Test B; AM= Austin Maze; TOH= Tower of Hanoi; BD = Block Design.
Ø = The potential score of this variable is not available due to post hoc data transformations.
∞ = No maximum score is set by the task.

8.2.3.2 Model Estimation and Specification

A CFA using MLE was performed using data from 102 participants. The independence model that tests the hypothesis that all variables are uncorrelated was rejected, $x^2(10, \mathit{103}) = 47.736$, p= <.001. The hypothesised model was then tested and demonstrated a good fit to the observed data, $x^2(5, \mathit{103}) = 6.72$, p= .242, SRMR = .053, RMSEA =.058, CFI = .954, GFI = .973. An evaluation of the standardised residuals did not reveal any aberrant values that would indicate poor fit (Appendix N).

As shown in in Table 36, MLE of the hypothesised one-factor model demonstrated significant relationships between all indicator variables and the COMPLEX & FAMILIAR latent factor.

Table 36

Maximum Likelihood Estimates of C&F GDC Model.

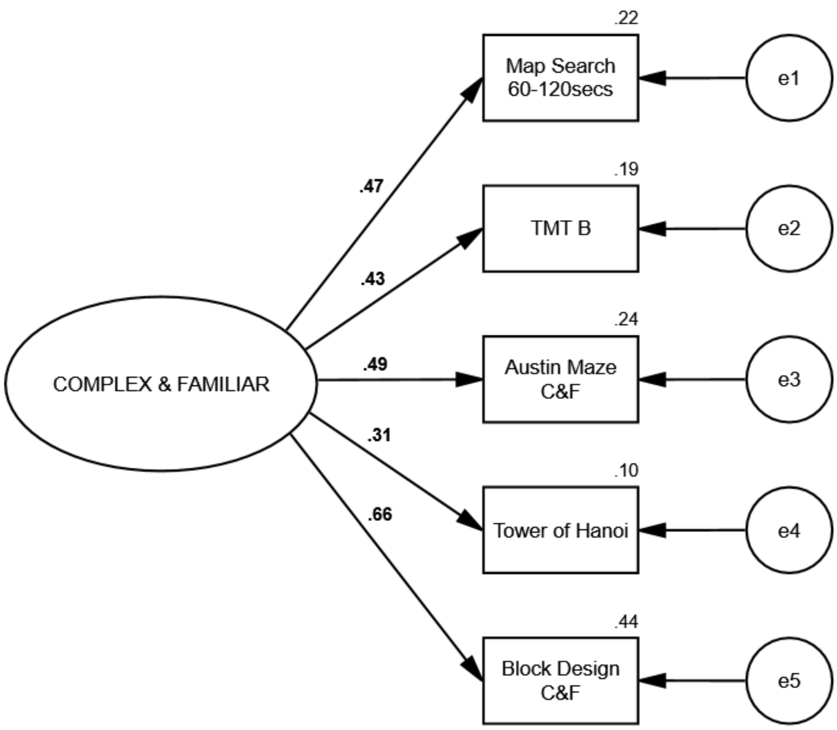
	Estimates				
Parameters	Standardised	Unstandardised	SE	Residual[a]	*p*
Map Search 60-120secs ← C&F	.467	3.916	1.038	1.811	<.001
TMT-B ← C&F	.431	6.783	1.942	0.531	<.001
AM C&F ← C&F	.486	1.433	0.366	2.510	<.001
TOH ← C&F	.308	0.820	0.329	8.541	.013
BD C&F ← C&F	.665	1.052	0.208	0.619	<.001

Note. TMT= Trail Making Test B; AM= Austin Maze; TOH= Tower of Hanoi; BD = Block Design; SE = Standard Error.
[a]Unstandardised residual error variance associated with each trial.

As seen in Figure 15, SMC values indicated that individual item reliabilities were moderately low overall. However, all variables were retained as the aim of this analysis was focused on the identification of a measurement model to determine correct factor weightings for the creation of a C&F composite score. Coefficient *H* was calculated to test the overall reliability of the model, which indicated an overall good reliability of the COMPLEX & FAMILAR model (0.632). The final model, including coefficients in their standardized form, is illustrated in Figure 15.

Figure 15

Final C&F GDC Model



Note. Austin Maze C&F = Austin Maze Complex & Familiar Score. Block Design C&F = Block Design Complex & Familiar Score. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each indicator variable are the squared multiple correlations.

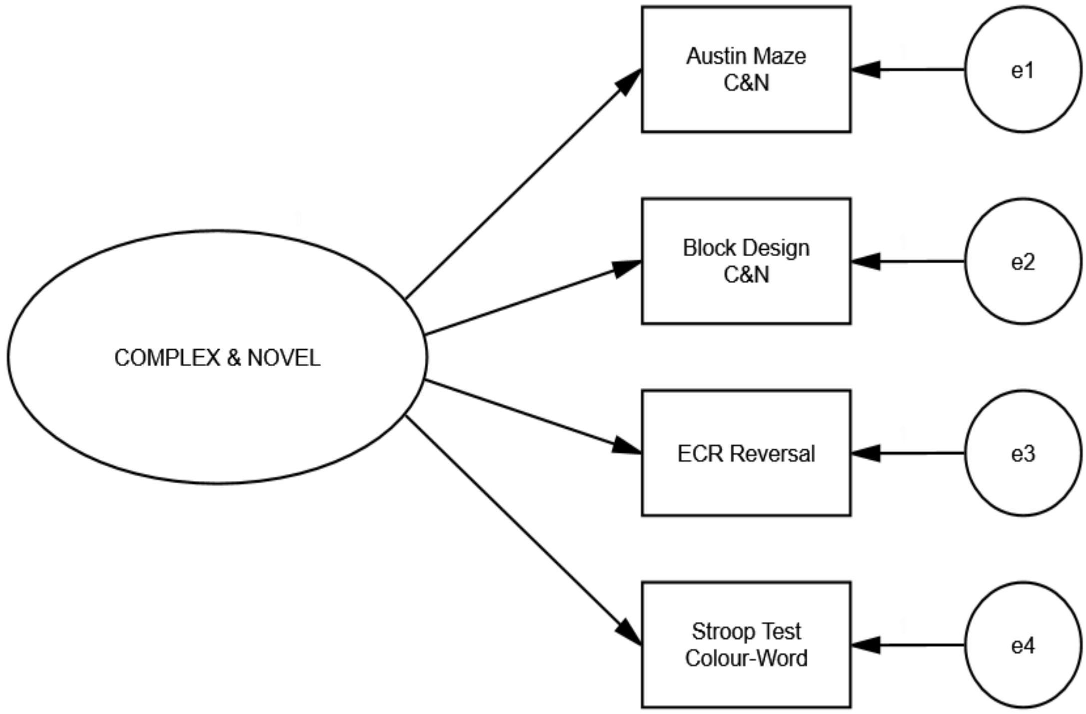
8.2.4 Complex and Novel

A CFA was performed using IBM SPSS AMOS v25.0 to investigate whether neuropsychological tests and task elements identified via the DCS to comprise of a C&N GDCs represented a 1-factor model solution (Figure 16).

.

Figure 16

Hypothesised C&N GDC Model



Note. Austin Maze C&N = Austin Maze Complex & Novel Score. Block Design C&N = Block Design Complex & Novel Score. ECR = Elevator Counting with Reversal. The larger circle represents the endogenous latent variable, squares represent exogenous indicator variables, and the small circles represent the residual variance.

8.2.4.1 Assumptions

As shown in Table 37, Skewness and Kurtosis values were within the acceptable range to support the assumption of univariate normality, and Mardia's estimate was within the acceptable range to support the assumption of multivariate normality. M-distance did not exceed the critical x^2 for df = 2 (α=.001) of 13.82 for any cases, indicating that multivariate outliers were not of concern. There were no missing data.

Table 37

Descriptive Statistics for C&N GDC Model Variables

Variable	*N*	*M*	*SD*	Range		Skewness	Kurtosis
				Potential	Actual		
AM C&N	103	25.98	3.29	Ø	15.54-30.95	-1.396	1.676
BD C&N	103	6.08	1.81	Ø	2.31-9.66	0.010	-0.900
ECR	103	7.82	2.23	0-10	2.00-10.00	-1.154	0.455
Stroop Test Colour-Word[a]	103	22.93	5.08	0 - ∞	9.00-33.90	-0.472	-0.122
Mardia's estimate							0.645
M-distance							13.57

Note. AM= Austin Maze; BD= Block Design; ECR = Elevator Counting Reversal
∞ = No maximum score is set by the task.
Ø = The potential score of this variable is not available due to post hoc data transformations.
[a]Values displayed are an inverse of the original score.

8.2.4.2 Model Estimation and Specification

A CFA using MLE was performed using data from 103 participants. The independence model that tests the hypothesis that all variables are uncorrelated was rejected, x^2(6, *103*) = 64.530, p=<.001. The hypothesised model was then tested and demonstrated an overall poor fit to the observed data, x^2(2, *103*) = 8.458, p= .015, SRMR = .0643, RMSEA = .178, CFI = .890, GFI = .963 AIC = 24.458.

Post hoc modifications were performed in an attempt to develop a better fitting and possibly more parsimonious model. Modification indices indicated that an improvement in model fit would be gained with the addition of a covariance between the residual error variances of *AM Complexity* and *Stroop Test Colour-Word*. A covariance path was added and the model re-estimated, which improved the overall model fit to the observed data, x^2(1, *103*)= .655, *p*= .418, SRMR = .0188, RMSEA = <.001, CFI = 1.00, GFI = .997, AIC = 18.655. A chi-square difference test indicated that not including this additional path would have significantly worsened the overall model fit, χ^2 diff (1, *n*=103) = 7.803, *p*= .005. AIC values also demonstrated improved parsimony of the model fit with the addition of the covariance path. As shown in in Table 38, MLE of the hypothesised one-factor model demonstrated significant relationships between all indicator variables and the latent factors.

Table 38

Maximum Likelihood Estimates of C&N GDC Model.

	Estimates				
Parameters	Standardised	Unstandardised	SE	Residual[a]	*p*
AM C&N	.567	1.858	0.389	7.306	<.001
BD C&N	.628	1.134	0.194	1.974	<.001
ECR	.679	1.509	0.242	2.665	<.001
Stroop Test: Colour-Word	.528	2.671	0.607	18.447	<.001
Residual Error Covariances					
AM C&N ↔ Stroop Test: Colour Word	-.365	-4.242	1.556		.006

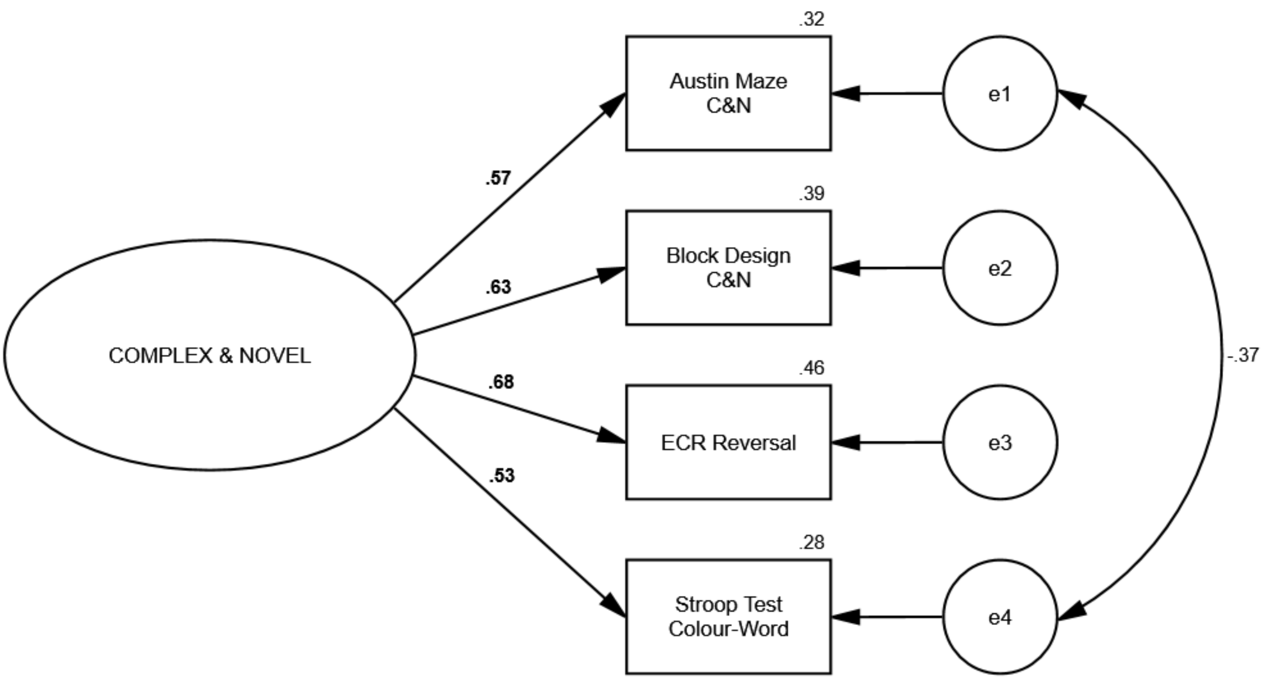
Note. AM= Austin Maze; BD= Block Design; ECR = Elevator Counting Reversal; SE = Standard Error.

[a] Unstandardised residual error variance associated with each trial.

As seen in Figure 17, SMC values indicated that individual item reliability was moderately low overall. However, all variables were retained as the aim of this analysis was focused on the identification of a measurement model to determine correct factor weightings for the creation of a C&N composite score. Coefficient *H* was calculated to test the overall reliability of the model, which indicated an overall good reliability of the COMPLEX & NOVEL model (0.70). The final model, including coefficients in their standardized form, is illustrated in Figure 17. In an effort to support readability, a summary of each GDC Model obtained during Study 2 and its corresponding variables is displayed in Figure 18.

Figure 17

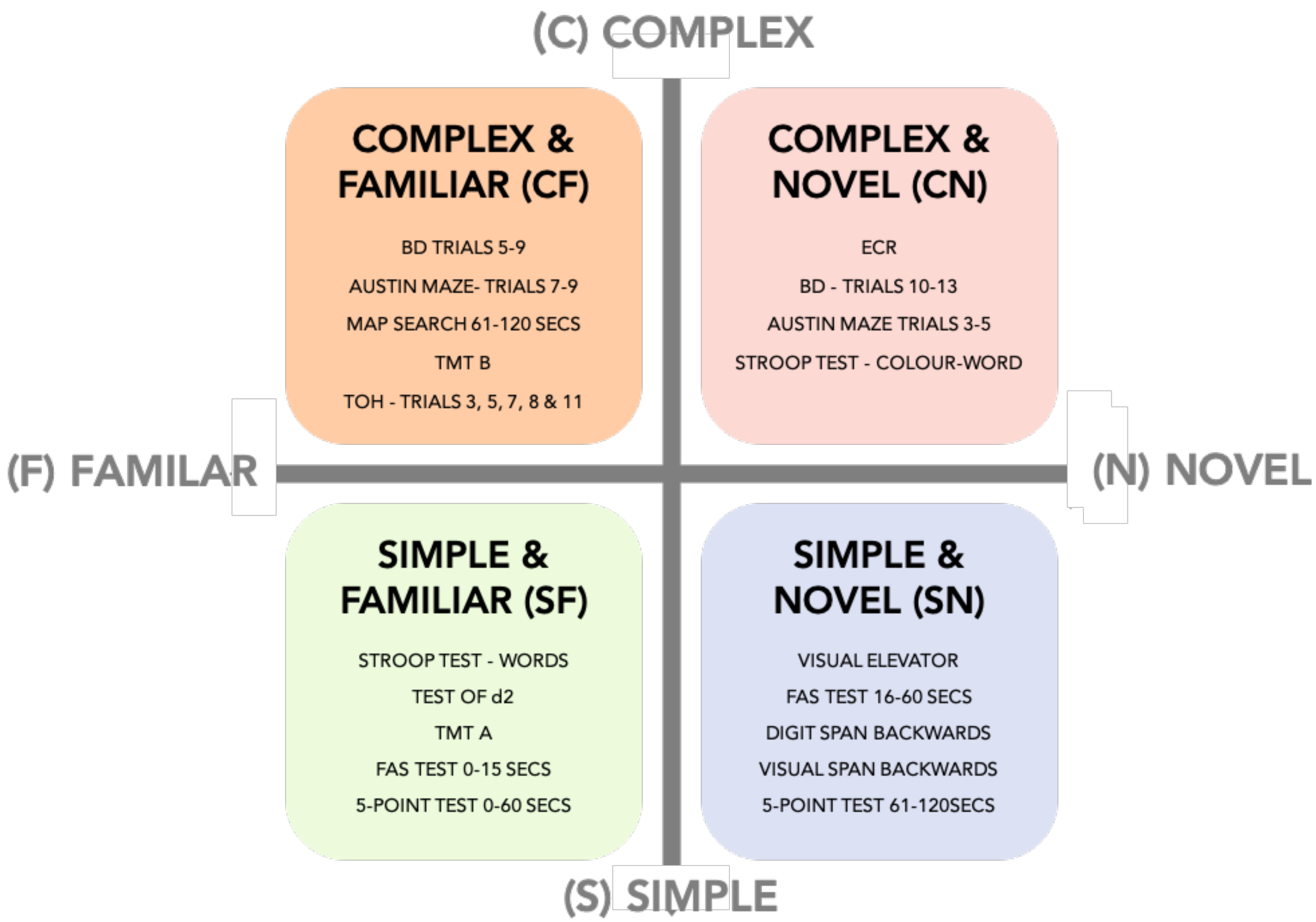
Final C&N GDC Model



Note. Austin Maze C&N = Austin Maze Complex & Novel Score. Block Design C&N = Block Design Complex & Novel Score. ECR = Elevator Counting with Reversal. The numbers superior to the each of the single headed arrows displayed in bold type face are the standardised factor loadings. The numbers above each indicator variable are the squared multiple correlations.

Figure 18

Summary of GDC Model Variables Represented Across a Dual Axis of Demand.



Note. BD = Block Design; TMT = Trail Making Test; TOH = Tower of Hanoi; ECR = Elevator Counting Reversal.

8.3 Discussion

Study 2 was necessary in order to establish the relationships between tests with similar GDCs from Study 1. Each of the four models demonstrated a good fit to the data and were accepted as measurement models of each GDC (S&F, S&N, C&F, C&N). The models were found to represent reliable GDC measures, with all demonstrating a significantly good fit to performance data. From a mathematical perspective, these findings strengthen the notion that cognitive control inherent in EF tasks may offer an additional explanation for performance differences. From the conceptual perspective, a prominent pattern of demand is revealed. As previously explained, each GDC is based on a varying combination of four Global Demands, (C) Complex, (S) Simple, (N) Novel, (F) Familiar. Each Global Demand is determined by criteria from distinct Demand Features (e.g. abstraction, contextual stability, etc.). Initial classification from Study 1, enforced the notion that complexity cannot be considered exclusively of novelty. That said, analysis of the details reveals a distinct pattern that can be identified across each axis in isolation, whereby (1) abstraction is the prominent Demand Feature of the complexity axis, and the novelty axis is weighted heavily by (6) automaticity on the novel end, and (7) schematic demands at the familiar end. These concepts will be further explored within this discussion.

8.3.1 The Significance of Abstraction within the Complexity Axis

Complex (1.C) abstraction demands were present in all tests represented by both complex GDC models (C&F, C&N). The prominence of (1) Abstraction demands across all GDC models supports previous research by Höchli et al. (2018) who proposed that abstraction may be one of the most fundamental characteristics of goal-directed behaviour. The current research demonstrates that the fundamental nature of abstraction coincides closely with the complexity of the overall task environment. (C) Complex Global Demands represented more sophisticated task requirements that included a set of subordinate goals within a complex superordinate goal state, requiring the evaluation of relationships and the forecasting and evaluation and correct sequencing of possible response actions (Höchli et al., 2018).

Both simple GDC models represented tasks that collectively comprise the most elementary form of abstraction requirements. Many of the tests that were represented required only the direct interaction and execution of subordinate goals that defined what and how to carry out a particular response action (Carver & Scheier, 2001). When contrasting the presence of (1.T) temporal, (1.R) relational and (1.P) policy abstraction demands, all were found to vary somewhat within and between each GDC model. Notably, S&F and S&N GDC models all shared the same (1P.S) simple policy abstraction demands. When considered alongside the (2.S) simple contextual stability that was also shared between the two, both S&F and S&N GDC models can be considered to collectively represent task environments that offer stable parameters, and that require minimal evaluation of task-specific relationships, policies, or reordering to formulate an appropriate response.

However, even though (1) abstraction demands were found consistently within simple and complex GDC models, a defining feature of complex GDCs was the variability in the type of (1) abstraction that was required. Unlike simple GDC models, where simple policy abstraction was prominent, complex GDC models showed more variable (1) abstraction requirements. This is consistent with previous proposals by Badre and Nee (2018), who suggested that many complex tasks require multiple forms of abstraction to complete. The complex GDC models developed for the current study demonstrate this, suggesting that a central indication to the complexity of a task may be demonstrated by variability in the forms of (1.C) Complex abstraction that are required.

Previous research has suggested that the level of abstraction demands may broadly depend on the structure of schemas held within a person's memory (Badre & Nee, 2018). This relationship was highlighted within the S&F GDC model in which (1.S) Simple abstraction demands and (7.F) Familiar schematic demands were found to consistently occur together across all test elements that were represented. A similar trend was also found for the C&N GDC model, with all but one test encompassing both (1.C) Complex abstraction and (7.N) Novel schematic demands. While this was not a completely shared demand, this trend suggests a relationship between (1) Abstraction and (7) Schematic Demands.

8.3.2 Automaticity and Schematic Demands are Polarised on The Novel Axis

(7.F) Familiar schematic demands were found to feature consistently within all tests that were represented by the familiar GDC models (S&F, and C&F). This indicates that (F) Familiar Global Demands represent an environment where the individual is able to have direct engagement with the task, with the ability to refer to and/or apply fundamental schemas to the task environment if needed. These demands may require the use of fundamental schemas surrounding language and semantics (FAS 0-15secs, Stroop Test-Word Trial. TMT-A) or shapes (BD C&F & 5-point 0-60secs).

Within some testing environments, (7.F) Familiar schematic demands were found to represent a task environment where (7.N) Novel schematic demands had already featured and had the opportunity become learned via exposure (e.g. AM C&F, TOH C&F). These requirements were demonstrated within the AM where similar (C) Complex Global Demands were found within the entire test, but later trials lessened in their (N) Novel Global Demands. This difference was attributed to learning of the hidden maze during (7.N) Novel schematic demands to form task-specific schemas which support the successful learning of the maze during later trials.

(6.N) Novel Automaticity was found to feature consistently within all tests that were represented by novel GDC models (C&N, and S&N). This indicates that (N) Novel Global Demands may largely reflect an environment where the formation of new task-specific behaviours beyond the use of implicit procedural knowledge is required. This task environment may require the manipulation of task-specific apparatus that is effortful (e.g. the manipulation of additional block apparatus during the BD C&N task), or required the inhibition of automaticity over a new task-specific behaviour (E.g. suppression of reading ability in favour or colour naming ability during the Stroop Test Colour-Word trial).

What this pattern of Demand Feature loadings represents, is that essentially, all C&N tasks will require (1.C) Complex abstraction and (7.N) Novel automaticity, amongst other demand features. All C&F tasks will require (1.C) Complex abstraction and (7.F) Familiar schematic Demands. Similarly, all S&N tasks will require (1.S) Simple abstraction and (7.N) Novel automaticity. Finally, all S&F tasks will require a (1.S) Simple level of abstraction and (7.F) Familiar schematic requirements. The significance of this pattern is not only that the dual axis of complexity and novelty are supported, but the defining demand features of these axes are identified at each end.

8.3.3 Conclusion

Study 2 demonstrated that performance can be captured between a set of EF tests in relation to their shared GDC. This allows for comparisons not only between tests but across demands that are shared between them, which offers a promising insight into how performance within neuropsychological measures may be further understood. However, while considerable similarities were observed within each GDC model, it is clear from the current study that the nature of demand is governed by its accumulation of contributing demand features, and not their mere addition. This natural variability was evident by the shared and unshared Demand Features within each GDC model that ultimately cluster to represent Global Demands across a dual axis. The findings presented here infer that the more S&F task demands are, the more distinguishable demand features within a test become, and thus the required responses are mostly direct and identifiable. However, Complex demands attract variability, and therefore the less pure and distinguishable the demands themselves are likely to be. Individuals approach solving complex problems from a variety of different perspectives, thereby creating a degree of “demand impurity”. To avoid perpetuating issues of ‘impurity’ amongst cognitive measures, the relationships between GDC performance models must be further explored. The current study demonstrated that the dual axis of demands is mostly and consistently, anchored by varying demands for abstraction, schemas and automaticity. Therefore, it is important to determine how their presence may be managed at various stages and represented by performance at various levels of demand.

Chapter 9

Study 3: Analysis of GDC Performance Relationships

Analyses performed during Studies 1 and 2 established four reliable GDC Models (S&F, S&N, C&F, C&N) that represented performance during various demand environments. Previous neuroscientific research has demonstrated the existence of a hierarchical rostro-caudal pattern of activation in response to the demands of the testing environment. The pattern of activation has been well documented to occur across a hierarchical gradient, with the upregulation of caudal regions and activation of rostral regions when demands for cognitive control increase. The structure of the DCS was informed by the consolidation of this research in order to capture performance in relation to increasing demand for cognitive control. Collectively, the criteria of the DCS has demonstrated the ability to identify (C) Complex and (N) Novel Demands where previous neuroscientific research has demonstrated the upregulation of rostral neural activity in response. The success of the DCS to distinguish task environment demands both within and between testing paradigms during Studies 1 and 2 provided a framework whereby performance during different levels of demand could be observed across a neuropsychological test battery. The aim of Study 3 was to explore whether a hierarchical relationship exists between performance outcomes of the four GDCs that reflects the graded rostro-caudal hierarchy of cognitive control in the brain. Specific hypotheses are drawn in the following section.

9.1 S&F → S&N

The S&F GDC model from Study 2 represented a collection of task environments that comprised of minimal (C) Complex and (N) Novel demands. Given that performance during S&F demands is largely reflective of the direct engagement with a testing environment that requires minimal cognitive control to coordinate a response, it was hypothesised that the S&F GDC Model would represent foundational performance within an overall hypothesised hierarchy of demands. S&F and S&N GDC models are differentiated due to the presence of (N) Novel demands. (N) Novel demands are considered indicative of the additional recruitment for cognitive control, largely due to the task environment demanding the formation of task-specific schemas and knowledge to formulate an appropriate response. As the formation of new schemas and knowledge arguably requires the need to initially recognise and access fundamental schemas prior to any adaptation, it was hypothesised that performance abilities during S&F demands would predict performance in response to Novel Demands.

9.2 S&N → C&F

The presence of (C) Complex demands is also indicative of an increase in cognitive control requirements. Study 2 demonstrated that (1) abstraction demands were found to combine with each GDC for (C) Complexity. Previous researchers have suggested that demands for abstraction may be influenced by the familiarity that is present in the testing environment. For example, when knowledge of task-specific rules or policies are available due to previous experience, the overall demand for abstraction may potentially be decreased. Moreover, when considered in the context of the dual axis of demand and GDC models, the ability to establish novel task-specific behaviours likely supports the ability to understand abstract complex relationships when the task environment is complex. Therefore, as the S&N model represented responses to task environment that required (1.S) Simple abstraction in order to formulate (N) Novel schemas and task-specific responses, it was hypothesised that performance during C&F demands would be predicted by performance during S&N Demands.

9.3 C&F → C&N

Due to the presence of both (C) Complex and (N) Novel demands within the same testing environment, performance during C&N demands is considered to represent a task environment that would require the most recruitment of cognitive control. As these demands arguably pose the most challenge in relation to all GDC conditions, performance here was considered to represent the apex of the hierarchy of demands. Therefore, the ability to successfully respond to C&N demands should theoretically be predicted by performance under lesser demands. Under the DCS, this would suggest that performance during C&N demands is supported by the ability to engage and respond to task environments that afford higher levels of simplicity or familiarity. Additionally, performance under C&F GDC demands should closely align due to the shared demands encountered for complexity. As the current project proposed that the ability to successfully negotiate (N) Novel demands is supported by the ability to first engage with known environments successfully, it was hypothesised that performance during C&N demands would be predicted by performance during C&F demands.

The hypotheses for Study 3 are summarised as follows:

1. S&N Demand Scores would be predicted by S&F Demand Scores,
2. C&F Demand Scores would be predicted by S&N Demand Scores, and
3. C&N Demand Scores would be predicted by C&F Demand Scores.

9.4 Method

9.4.1 Structural Equation Modelling

Structural Equation Modelling (SEM) was suited to the nature of the hypothesis testing (confirmatory) approach to of the current study by enabling the testing of structural theory underpinning performance data (Byrne, 2016). SEM is able to measure the causal processes under investigation via a series of regression equations between latent and observed variables. The advantage of SEM over other multivariate analyses is the capacity to run a simultaneous analysis of the entire set of variables (both latent and observed) of a hypothesised model to measure the extent to which the structure of the model is consistent (fit) with the observed data. Evaluation of whether the hypothesised model demonstrates a ‘good’ fit is further informed via the statistical fit indices in Section 5.5.4.

9.4.2 Measurement Model and Structural Models

Fundamentally, a Structural Equation Model contains both a measurement portion and a structural portion. The measurement portion is concerned with the psychometrically appropriate measurement of latent factors via observed indicator variables, whereas the structural portion is concerned the structural relationships between the latent factors. Studies 1b and 2 established a measurement model of the GDCs whereby structural relationships can be assessed during Study 3. Seminal authors have proposed that best practice for SEM should follow this two-step approach whereby; (1) the measurement portion of the hypothesised model are tested prior to the (2) structural relationships between the latent factors are tested (Anderson & Gerbing, 1988, 1992)

9.4.3 Full Structural Model Specification

In order to appropriately specify the full structural model, weighted composite scores that represented each GDC model in Study 2 served as the observed variables in Study 3. This approach enabled the interpretation of performance in regard to each GDC. Accordingly, each GDC Score provides an appropriate representation of performance during four different environmental demands (S&F, S&N, C&F, C&N). The calculation of parameter estimates using composite scores derived from single one-factor congeneric models enabled a more stable analysis in comparison to a single model containing multiple original indicator variables that converge onto the hypothesised latent factors. This approach enabled the maximum number of available cases per parameter estimate to meet SEM sample size assumptions within the current study to be fulfilled.

Coefficient *H* values for each composite score were also fixed in the final structural model to provide specification of the amount of error associated with the measurement of each latent variable. By calculating coefficient *H* prior to the estimation of the full structural model, the regression coefficients for each composite indictor variable and error variances were then able to be calculated and fixed to provide further stability to the SEM analysis (Levine et al., 2005; Munck, 1979). Regression coefficients for each composite score were calculated using the Munck (1979) formula as follows:

$$\lambda = \sigma(\mathrm{x})\cdot\sqrt{\mathrm{r}}$$

Note: λ = is the regression coefficient, $\sigma(x)$= is the standard deviation of the composite variable, r = is the reliability of the composite variable

Furthermore, Error variances for each composite score were calculated via the following formula:

$$\theta = \sigma^2(\mathrm{x})(1\text{-}\mathrm{r})$$

Note. θ = error variances, $\sigma^2(x)$= is the variance of the composite variable, and r= is the reliability of the composite variable

9.4.4 Procedure

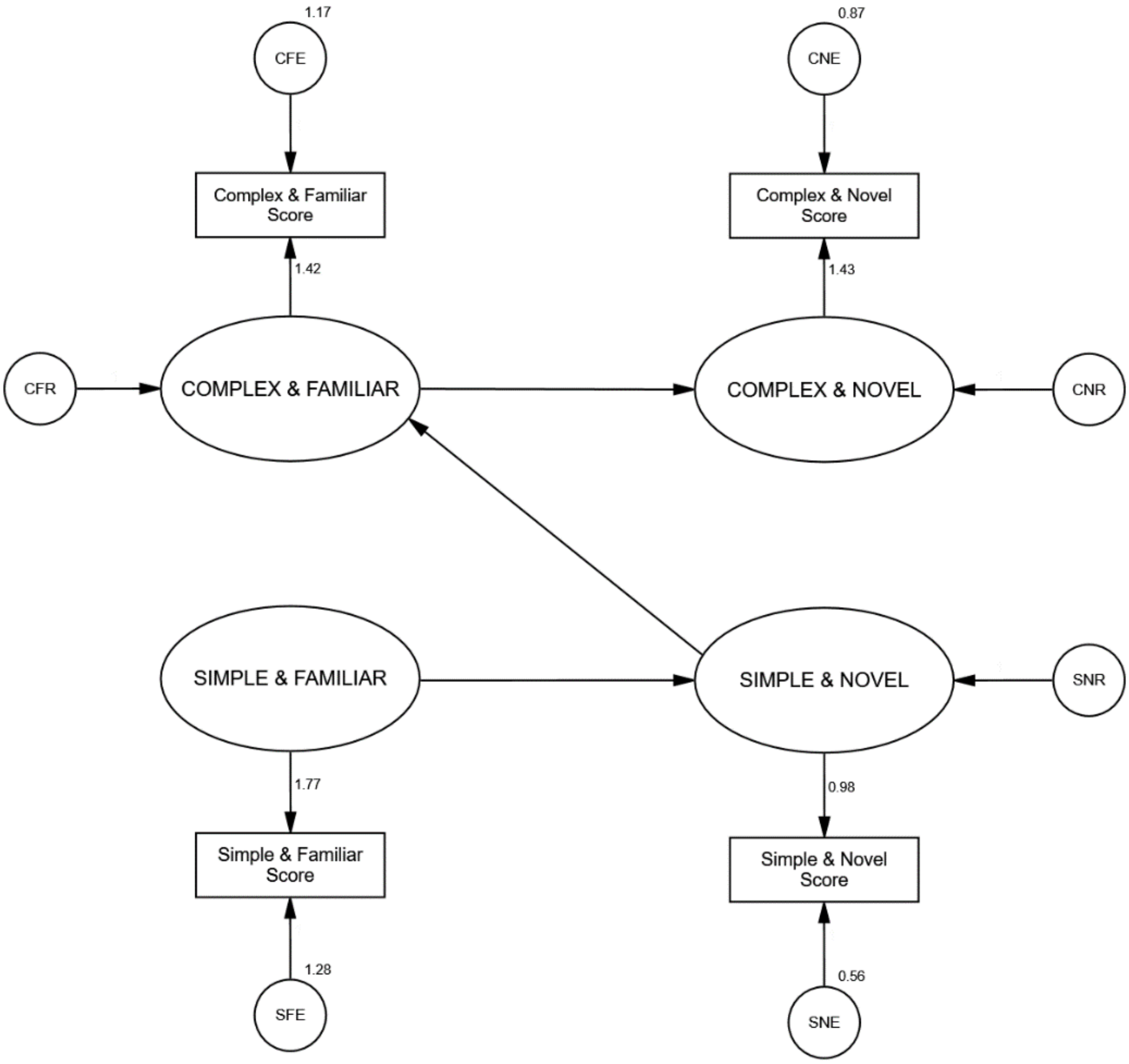
GDC scores were entered into an SEM model as indicator variables with their known weighted factor loading (Appendix P) and error variances assigned. This resulted in establishment of a structural model whereby latent factors were representative of GDC performance scores for which unique performance variance was empirically known, and associated error variance was controlled.

An autoregressive quasi-simplex model (QSM) was hypothesised to represent the hierarchically contingent set of relationships between the GDC performance scores. The QSM enabled the set of latent variables to be ordered on a unidimensional scale, where measures closest to the scale are highly correlated. For the model in Figure 19, it was hypothesised that an autoregressive process would exist between the latent factors that represented each GDC

Score. This proposes that each factor is significantly regressed on the previous factor in the path sequence. This means that the first GDC Score at the start of the path sequence in the model should occur prior to the next in the sequence, and second GDC Score occurs before the third GDC Score in the path sequence, and so on. The autoregressive nature of the model meant that relationships could be assessed in regard to the ability of each GDC Score to predict future performance based on past performance. As the nature of the scoring is not based on time-scale data, the relationships reflect a sequential model of performance scores from S&F demands through to C&N demands.

Figure 19

Hypothesised SEM of GDC Performance Score Including Known Factor Loading



Note. SFE = Simple & Familiar Error Variance; SNE = Simple & Novel Error Variance; CFE = Complex & Familiar Error Variance; CNE = Complex & Novel Error Variance. SNR =Simple & Novel Residual Variance; CFR = Complex & Familiar Residual Variance; CNE = Complex & Novel Residual Variance.

9.5 Results

9.5.1 Assumptions

As shown in Table 39, Skewness and Kurtosis values fell within the acceptable range to support the assumption of univariate normality, and Mardia's estimate indicated no potential violation of multivariate normality. Analyses of M-distance revealed two cases (M-Distance = 16.15 and 14.29) that exceeded the critical x^2 for df = 2 (α= .001) of 13.82. In light of the current sample size being able to meet conventional requirements for SEM analyses with the removal of these cases, the cases were excluded from the analysis. Re-computation of M-distance with n= 100 revealed no cases that exceeded the critical x^2 for df = 2 (α = .001), indicating that multivariate outliers were no longer of concern. Finally, collinearity diagnostics were performed with COMPLEX & NOVEL as the predicator variable and reported acceptable tolerances and VIF (Variance Inflation Factor), indicating that multicollinearity would not interfere with the interpretation of the analyses. There were no missing data.

Table 39

Descriptive Statistics of all GDC Scores.

Variable	*N*	*M*	*SD*	Range	Skewness	Kurtosis	Tolerance	VIF
S&F Score	100	10.9	2.11	4.62 – 14.59	-0.565	0.033	.456	2.195
S&N Score	100	5.76	1.24	2.80-8.27	-0.128	-0.500	.467	2.139
C&F Score	100	15.23	1.73	10.79- 18.42	-0.359	-0.444	.831	1.203
C&N Score	100	10.57	1.59	6.36-13.42	-0.678	-0.082	-	-
Mardia's estimate								-.979
M- Distance								11.00

Note. S&F = Simple & Familiar; S&N= Simple & Novel; C&F= Complex & Familiar; C&N= Complex & Novel; VIF= Variance Inflation Factor.

As observed in Table 40, the correlation matrix demonstrated a trend of a simplex pattern amongst the variables. A simplex pattern is such that the correlation values begin to decrease as one moves away from the diagonal values. Thus, the hypothesised structural relationships of the model were considered suitable for further analyses.

Table 40

Pearson's Correlation Matrix of GDC Scores

	S&F Score	S&N Score	C&F Score	C&N Score
S&F Score	1.000			
S&N Score	.725	1.000		
C&F Score	.394	.365	1.000	
C&N Score	.534	.501	.798	1.000

Note. S&F = Simple & Familiar; S&N= Simple & Novel; C&F= Complex & Familiar; C&N= Complex & Novel.

9.5.2 Model Estimation

SEM using MLE was performed using data from 100 participants. The hypothesised model demonstrated a good fit between the model and the observed data, $x^2(2, \mathit{100}) = .250$, $p = .882$, SRMR = .0075, RMSEA = <.001, CFI =1.00, GFI =.999, AIC= 16.250. An evaluation of the standardised residuals did not reveal any aberrant values that would indicate poor fit (Appendix Q). As seen in Table 41, all path estimates reported a significant relationship. It was observed that within the standardised estimates, two values exceeded '1'. This was not of concern as SEM factor loadings are regression coefficients and not correlations where standardised values larger than '1' in magnitude are possible (Jöreskog & Sörbom, 1989). Modification indices were requested via IBM SPSS AMOS V25.0 to explore whether any potential mediating regression paths that were not specified within the hypothesised model were detected. No additional paths were identified; thus, the current model was accepted. The final model can be observed in Figure 20.

Table 41

Maximum Likelihood Estimates for Full Specified GDC Model.

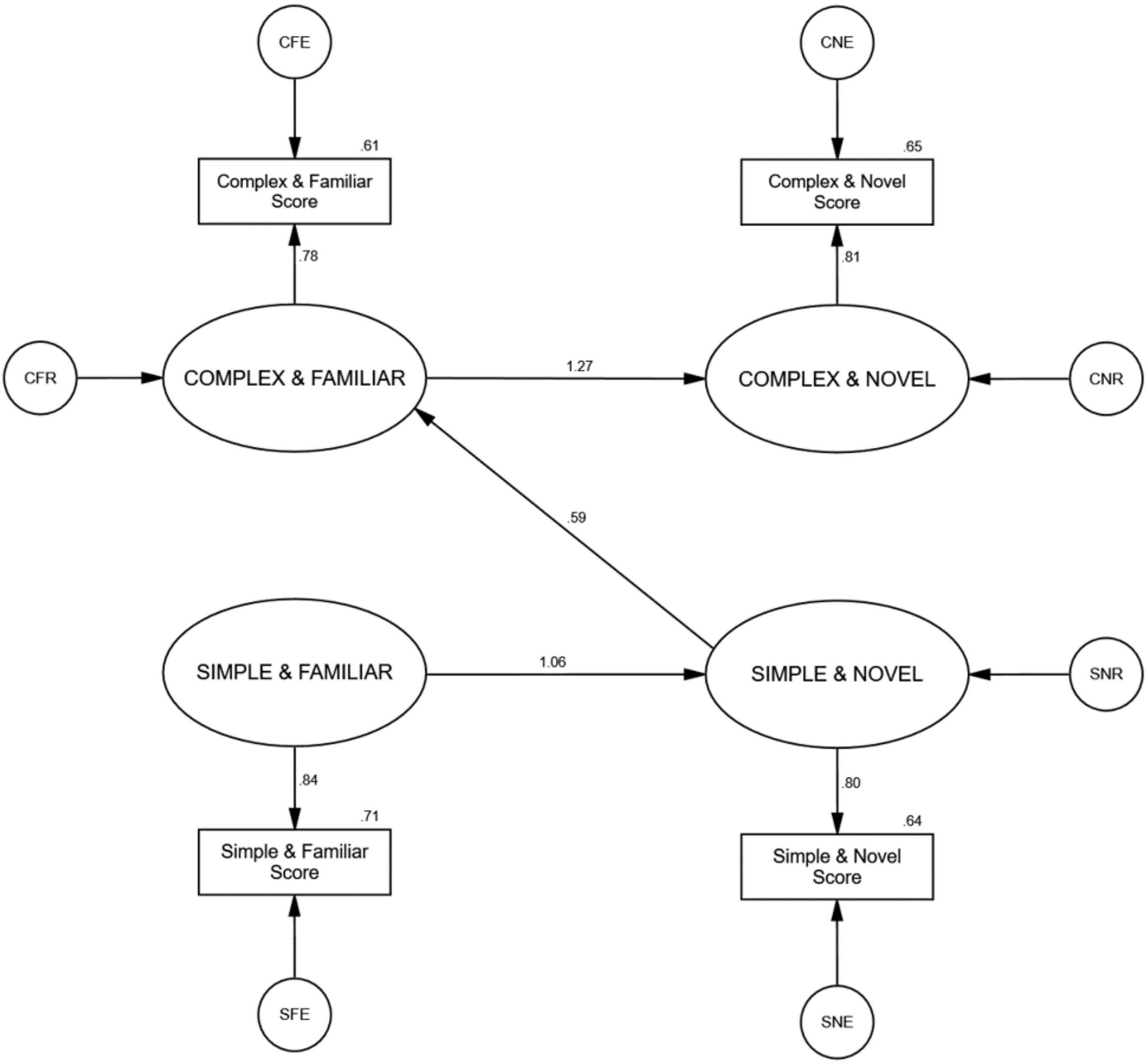
	Estimates				
Parameters	Standardised	Unstandardised	SE	Residual [a]	*p*
S&N ← S&F	1.062	1.074	.121	-0.132	<.001
C&F← S&N	0.579	0.532	.125	0.578	<.001
C&N←C&F	1.289	1.236	.183	-0.530	<.001

Note. SE= Standard Error

[a] Unstandardised residual error variance associated with each trial.

Figure 20

Final SEM of GDC Performance Scores



Note. The numbers superior and adjacent to the each of the single headed arrows are the standardised path coefficients. The numbers above each indicator variable are the squared multiple correlations.

9.6 Discussion

Study 3 of the project aimed to explore whether a hierarchical relationship existed between performance outcomes of GDC performance scores, as relationships represented are consistent with the purported graded rostro-caudal hierarchy of cognitive control in the brain. An autoregressive QSM was hypothesised that depicted dependent relationships between four GDC Scores. A model was accepted, and the hypotheses supported for relationships that collectively demonstrated a hierarchical pattern of performance during various GDCs.

9.6.1 S&F → S&N

The hypothesis that the S&N GDC scores would be predicted by the S&F GDC score was supported. A large coefficient was reported between these two GDC scores that indicated a strong shared relationship. While the strength of this relationship to an extent reflects the consistent (S) Simple demands shared between the S&F and S&N environments, when conceptualised across a dual axis it also signifies that successful performance during (N) Novel demands is contingent on performance during (F) Familiar demands, when the complexity of the task environment is held constant. The strength of the relationship between S&F and S&N demands is not surprising, given it reflects the shift from (F) Familiar to (N) Novel demands, essentially a shift in the horizontal axis.

The current findings support the proposal of previous researchers (Bhandari & Badre, 2018; Cole et al., 2011), whereby the successful adaptation to (N) Novel demands is considered to rely on the knowledge from prior experience with similar tasks. S&F performance scores largely represent performance ability under (7.F) Familiar schematic demands, as indicated in Study 2. The dependent relationship found in Study 3 therefore further demonstrates a parallel to previous findings whereby performance under (N) Novel demands can be facilitated when sufficient representations are available that ultimately reduce the amount of additional learning that is needed (Cole et al., 2011; Collins & Frank, 2013). Therefore, the more familiar the stimulus and response representations that are stored in memory, the greater the influence on success during the formation of new task-specific representations when novel demands are present.

9.6.2 S&N → C&F

The hypothesis that C&F performance scores would be predicted by S&N scores was also supported, with a significant moderate positive path coefficient reported. Interestingly, no DCS Demand Criteria were consistently shared between the S&N and C&F GDC Models reported in Study 2. Both scores represented performance that occurs during one Global Demand that requires high levels of cognitive control; (C) Complex, or (N) Novel. This adds support that GDC scores are representative of performance across accumulated demands, and not merely capitalising on the relationships between similar Demand Features.

As abstraction demands are found to coalesce the complexity of each GDC Score, this relationship suggests that the success during (1.S) Simple abstraction demands when relationships, ordering of responses and overall responses are explicit, will support the ability to respond when relationships, and/or ordering of responses are required to be determined by the individual. How this relationship may be supported may be answered by previous research suggesting that the complexity of abstraction demands can broadly depend on the structure of information that schemas hold within a person's memory stores (Badre & Nee, 2018). This is also demonstrated within the current relationship, whereby (7.F) Familiar schematic demands within a C&F task environment may moderate effects of (1.C) Complex abstraction demands, thus supporting performance and potentially reducing the overall level of demand for cognitive control.

9.6.3 C&F → C&N

The hypothesis that C&N GDC scores would be predicted by C&F GDC Scores was also supported. This relationship is comparable to that found between S&F and S&N GDC Scores whereby (N) Novel performance may be supported by schematic representations that are held by the individual. During (C) Complex Global Demands, the ability to navigate both (N) Novel and (F) Familiar demands may be required within the same overall test administration. Study 1b of the project identified that two Global Demand conditions represented performance within the AM. These were attributed to the (N) Novelty and (F) Familiarity of the task demands across consistent (C) Complex demands of the task. During the AM, (7.N) task-specific schemas are encountered and required to be formed, in addition to the retention of (8.N) Novel episodic information. Successful performance during the AM is measured by how many error-free moves are made, inferring that the hidden maze has been learned and the prescribed rules have been followed. In order to achieve this, task-specific schemas are required to be formulated and stored to allow for repetition of the responses required from the initial engagement with (N) Novel demands. Therefore, performance within the AM is considered dependent on how efficiently and accurately the individual can perform during (F) Familiar demand trials, after overcoming the initial (N) Novel demands of the task. Failure to perform under (F) Familiar demands would suggest that (N) Novel Demands remains high, due to the task-specific schemas not being acquired. This suggests that under (C) Complex demands, the relationship reported by the model reflects that performance can be supported by how well the experiences during (N) Novel demands are converted, stored, and transformed to become familiar.

9.6.4 GDC Performance across a Hierarchy of Demand

The statistical modelling of GDC scores demonstrated a series of dependent relationships of performance under four distinctive, but interrelated, task environment demands. The GDC collectively represent a hierarchical pattern of performance during varied demands of complexity and novelty that would call for the recruitment of cognitive control. When compared to the neurological evidence that informed the structure and development of this project, similarities can be drawn between both the theoretical and neurological accounts of cognitive control mechanisms. Amongst neurological accounts of cognitive control, the current study may provide a quantified account, and support for, the graded nature of the EMN activation (Hugdahl et al., 2015). The hierarchical relationship between performance under varied demands within this study, when considered together with previous findings by Hugdahl et al. (2015), collectively highlight that foundational cognitive abilities can influence and support performance when demands for cognitive control increase. This performance hierarchy arguably reflects a system whereby tasks that require high levels of cognitive control resources (C&N) are dependent on the support and integration of resources that are engaged when demands are more singular, and less dependent on cognitive control resources in comparison (S&F, S&N, C&F). Hugdahl et al. (2015) observed that upregulation of cognitive control frontal networks when task demands increased to require cognitive control were in addition to already active caudal networks that support memory, orientation and motor actions.

As outlined in the current study, the quantified nature of the demands that govern this hierarchy should be considered complementary to previous control research regarding the frontal lobe activations that subserve responses to C&N demands. The frontal lobes are considered to be central to the selection and execution of top-down control of posterior cortical regions that are necessary to coordinate action when demands for controlled behaviour are high (Fuster, 2000a, 2002, 2017). The current model proposes a performance-based measure of this relationship, whereby the responses to C&N demands are contingent on the coordination of these foundational systems to support the response across all GDC levels. Across this hierarchy, the cognitive skills required to respond to S&F, S&N demands are considered predominantly singular and specifically measurable, whereas this specificity becomes more enigmatic at higher levels of this hierarchy when demands (C&F and C&N) increase, which require control systems to orchestrate and integrate these distinct cognitive skills to respond. Badre & Nee (2018) proposed that the most caudal subdivisions of the frontal cortex (Pre-motor cortex) support execution of sensory-motor control. Within the current study, S&F demand performance is most suitably aligned with the demands for caudal activation due to the singular engagement with schematically familiar sensory-motor S→R relationships.

The relationships between (N) Novel and (F) Familiar demand performance found within the current study likely represent the processes of the rostro-caudal gradient of activation. The integration of these seems relatively dependent on the automaticity that can be afforded by the task environment (Jeon & Friederici, 2015). During (N) Novel demands, performance is directed by the frontal lobes that exert top-down control over caudal regions to appraise and activate any applicable S→R knowledge (Badre & Nee, 2018; Fuster, 2000a, 2017). Frontal systems then work on integrating any S→R knowledge (schemas) into a series of new explicit task-specific behaviours in response to (N) Novel demands. Thus, performance under novel demands that reflect abstraction and schematic abilities likely reflects the efficiency of these frontal systems to coordinate foundational caudal systems in service to the establishment of a novel controlled response.

9.6.5 Conclusion

Study 3 of the project demonstrated that when demands for cognitive control are taken into account, performance that represents functional neurological recruitment can be captured across a test battery. Within these testing environments, a hierarchical pattern of performance exists that is dependent on gradients of complexity and novelty. Importantly, performance during higher gradients of demand are ultimately supported by cognitive control systems that subserve performance during lower levels of demand. The hierarchical relationships modelled here arguably provide a quantifiable account of how the gradient of rostro-caudal organisation can be attributed to task performance across a neuropsychological test battery.

The ability to discern a hierarchy of demand within a test battery has important implications. With this knowledge, the research collective can begin to approach test administration and interpretation with an additional understanding of the influential contribution that cognitive control, as operationalised by complexity and novelty, can have on performance. Further, interrelationships suggest that the traditional notion of targeting only C&N demands during assessment may be short-sighted if a full account of how this performance is managed is to be understood. In order to provide a more fulsome account, performance must be captured across intersecting continuums of demand in order to better isolate how performance is managed. The ability to demarcate this performance provides a greater understanding towards how responses are required to be executed and further elucidates *where* changes in this hierarchy can influence *how* the individual is required to respond.

Chapter 10

General Discussion

At the conceptual level the similarities between the constructs of EF and Cognitive Control are significant. However, attempts to operationalise and classify the constituents of either construct have highlighted that meaningful differences in the underlying mechanisms or components are apparent. Given that current EF conceptualisations remain mired in past notions that one task measures a singular cognitive component, advancing cognitive theory has remained elusive. Unsophisticated administration and scoring procedures perpetuate the notion of a poorly defined homunculus. Yet, while cognitive control theory recognises and prescribes a multifaceted system, there is an absence of any direct measures of its contribution to behaviour outside of the neuroimaging scanner. This thesis has attempted to take a novel perspective that is distanced from the typical approach of EF or cognitive control studies, by a) operationalising cognitive control demands b) reconceptualising scoring of EF tasks and c) assessing newly conceptualised performance outcomes against newly developed criteria of cognitive control. This approach argues that including a demand appraisal of tasks is critical to understanding how performance must be managed cognitively. This meaningful addition to cognitive assessment interpretation is likely not confined to, but is rather exemplified by, EF. In sum, the findings from the studies conducted include the following notable contributions to knowledge:

(a) Demands for complexity and novelty are quantifiable when the DCS is used to appraise how responses to neuropsychological test paradigms are applied in order to reach a successful outcome.

(b) Performance within a single neuropsychological test can vary significantly in response to changing demands, particularly when changes to the demand for novelty occur. This reinforces that a single outcome score is insufficient in the assessment of multifaceted performance demands.

(c) Performance outcomes from neuropsychological tests traditionally viewed as divergent in their purpose can represent significant communality when the demands for complexity and novelty are shared.

(d) Performance at various levels of demand reflects a hierarchically contingent continuum that is responsive to the gradient of complexity and novelty within a test environment.

The outcomes of this project hold significant importance for enhancing the ecological validity of tasks of known frontal lobe activation. Moreover, this contemporary approach that is offered may also serve to provide an alternative lens through which to view many of the pre-existing longstanding challenges of task impurity, and the classification of EF within cognitive theory. A central narrative of this project was the need to provide a framework for the measurement and interpretation complex human behaviour that has the potential to provide improved clinical utility EF tests were chosen as their environment requires cognitive adaptability to achieve goal-directed behaviour. Their inclusion and exploration via the methods developed by this project has demonstrated that the nature of this adaptive, goal-directed task performance is responsive to a hierarchy of demands. This has significant implications not only for the future use of these tools as singular measures of EF, but also how these measures can be better utilised for the assessment of overall higher-order cognitive ability. This approach urges the recognition of the synergy between cognitive control and EF, and the duality of their influence over the execution of controlled behaviour in response to demand. This project not only offers a methodical system whereby this synergistic approach can be successfully applied, but also a framework that is able to account for the nature of human engagement at various levels of novelty and complexity across a hierarchical continuum of demand.

10.1 Reconciling Theoretical Accounts of EF

It is evident from the current project that novelty within the goal-directed testing environment represents the application, update or transformation of schemas to task-specific controlled responses that are prescribed by complexity. In their current traditional administration, the human capacity to navigate this process is likely what many neuropsychological test performance outcomes are measuring, yet often fail to capture. Thus, in this context, the term novelty should be considered to encompass the degree of demand that is required, and not simply the overall ability of a test to present a 'new' environment. The recognition and integration of familiar and novel information is fundamental to human behaviour, as without it, each complex rule and action would need to be determined from the beginning each time (Cole et al., 2011). This process is well documented within the SAS model (Stuss et al., 1995). However, reconciling how this novelty is managed in the context of various demands for complexity has been notably elusive. The demarcation of these demands via the quantified structure provided by the DCS offers utility towards how any success and failures can be evaluated in respect to the individual management of novelty.

Similar to the tests that serve to inform them, contemporary EF theories are limited by their lack of specificity regarding the context or environment where their skillsets are deployed. The explanation of complex, goal-directed behaviour underpins the definitions of these theories, yet little attention is paid to positioning them in respect of the naturally diverse environments where these behaviours will undoubtedly be required. The nature of EF assumes adaptability in response to change, but fails to identify, demarcate and capture this in assessment practices. Many EF theories have alluded to an 'extra' component of the EF system that exists beyond just a singular skill set. For Miyake and colleagues (2012), this was attributed to a common EF factor that is present across all EF abilities they studied. For Anderson et al. (2011), this was arguably underpinning the nature of the relationships between each of the domains proposed.

The problem is that such EF models are categorical, and not continuum based. It has become ingrained within the neuropsychological literature to seek out and fragment systems to achieve purity in the provision of quantified instruments. This categorical approach has led to detailed accounts of key facets of EF that ultimately fail to sufficiently acknowledge the architecture of the system to which it belongs, and the consequent generation of conceptual homunculi and measurement impurity. This continued drive for extreme parsimony threatens to over-simplify the sophistication and variability of individual human behaviour. Decades of literature has been produced that began with process-driven accounts servicing 'how' different cognitive events occur, to the now contemporary domain centred accounts that prescribe the 'what' components. This focal level of analysis has served to provide deterministic accounts of behaviour, but theory attempting to bridge between the *what* and *how* concepts to account for the context of the variability within a test environment remains elusive.

The premise of the current project follows a similar line of investigation to the contemporary domain centred models by the development of a quantified Global Demand Classification system. However, the notion of purity is potentially false due to the very nature of striving to seek it, as modern accounts of neurological function have abandoned concepts of singular regional activation of neural structures. The current project was operationalised using cognitive control theory due to its contemporary recognition of a gradient of control that aligns with a natural environment, and not a purely categorical decision-making process (Badre, 2008; Badre & D'Esposito, 2009; Duncan, 2013; Koechlin, 2016). By doing so, the GDC seeks to provide an account of 'what' diversity exists within test environments, and analysis of the relationships between the different classifications seeks to provide an account of 'how' they contribute to previous notions of 'impurity'. However, the GDC for a test is not designed to replace or absolve the interpretation of EF skills, instead, the GDCs provide an additional account of performance variance that is attributable in the context that EF may be required. Thus, to provide a more complete account of performance, the GDC and knowledge of the EF skillsets must be applied in tandem to bring greater clarity to the assessment and analysis of goal-directed task completion. In doing so, the information that is provided by EF tests should be able to be assimilated into a unified approach to assessment whereby theoretical understandings of complex higher-order cognition can become reconciled.

10.1.1 Improving Ecological Validity by Identifying Demand

EF is arguably required depending on the goal requirements (e.g. form and deliver a plan, inhibit a response, update a sequence). When constructing and interpreting a test environment through the lens of EF assessment only, the most complex EF tests fail the basic elements of psychometric properties. Measures are reliable in that they repeatedly measure the same outcome, however, the validity of what they are measuring must be questioned. Issues with validity emerge since current EF theory is necessarily built on the foundation of EF test performance, and therefore subscribes to the notion that singular outcome scores sufficiently equate to performance. Instead, the current project has demonstrated that within many EF tests, multifaceted performance coincides with changes that require the individual to recognise, appraise and integrate information across a continuum of demand. The results of the current project argue that it is not the failure of the tests themselves to capture behaviour, but a failure in their interpretation. There is a fundamental need to evolve our approach to concepts of complex higher-order cognition and its assessment that extends beyond the constraints of EF. This project has shed light on the importance of this and provided a quantified account of the influence of performance in reference to demands for complexity and novelty, and thereby establishing a first step towards the coalescence of cognitive control and EF theories.

This project operationalised the dual axis of demand using the GDC. Study 1 demonstrated that performance *within* a multifaceted test (that is usually measured with a singular aggregate score) is not always equal between trials, and that this lack of equivalence between can be captured using the GDC. Thus, while a test may be purported as a measure of 'inhibition', inhibition abilities may be tested under several different GDCs, which ultimately influences performance outcomes. While the use of an aggregate score may arguably provide an 'average' measure of inhibition across these different environments, the current study demonstrated that when the demand for cognitive control is considered, a superior account of performance can be achieved.

If the interest of research is to explore particular skill sets, this must also occur across a variety of demands. For example, the Visual Span - Backwards test and the ECR test are both documented to require WM abilities for successful completion (Lezak et al., 2012; Nimmo-Smith et al., 1994; Strauss et al., 2006). Thus, performance scores on both tests are utilised to infer WM ability. However, the current study demonstrated that these two tests diverged in their Global Demands for complexity but shared the same Global Demands for novelty. It can therefore be concluded that these tests do not represent performance of WM abilities under the same environmental demands. By incorporating GDC knowledge from each test, it can be seen that the Visual Span-Backwards test represents working memory abilities within a Simple and Novel task environment, whereas performance on the ECR test represents working memory abilities within a Complex and Novel task environment. An acknowledgement of this difference is important in light of evidence from this project that demonstrates performance is influenced by the GDC of a test. Therefore, the evaluation of performance between tests that require similar cognitive skills in conjunction with their respective GDC may offer an understanding towards strengths and weaknesses in performance outcomes in relation to demand that may not be available when considering performance of a cognitive skillset in isolation. This increased understanding in relation to demand will strengthen the ecological validity of outcomes that can be drawn from the assessment of complex cognition, including EF.

10.2 The Hierarchical Demand Model (HDM).

The success of the GDC to capture differences in performance within a testing environment calls for the formation of theory that encapsulates changing demand. To this end, this thesis offers the Hierarchical Demand Model (HDM). The HDM proposes that in order to engage in a goal-directed task successfully, the individual must *recognise* what is known, *appraise* what is required*,* and *reconcile* the difference in order to formulate an effective response. The magnitude of difference between recognition of what is known, and the appraisal of what is required, will dictate the classifications of complexity and novelty. The HDM proposes that this process is influenced by, and ultimately a product of the environment being attributable to one of four GDCs; Simple & Familiar, Simple & Novel, Complex & Familiar, or Complex & Novel. Within the model, each of these GDCs serves to provide a quantifiable marker for the level of demand for adaptive behaviour that is engaged when complex behaviour is undertaken.

When faced with a goal-directed task, the individual must first *recognise* the features that are present within the environment to determine elements that bear any familiarity to what is already known. Here, the individual will search for familiar patterns to reference against stored knowledge and schemas. This may require the assimilation or accommodation of schemas, which is not unique to cognitive control or EF, but serves to support their engagement. The process of schema activation and consolidation has been put forward in previous research (e.g. Baddeley, 1998; Stuss et al., 1995). The extent of familiarity recognised in the stimulus ultimately affords a reprieve from demands of complexity and novelty and reduces the amount of cognitive effort that is exerted. Arguably, the efficiency of human cognition is owed to this process, whereby choice and selection is preferenced towards environments that are familiar and well-learned (Liao et al., 2011).

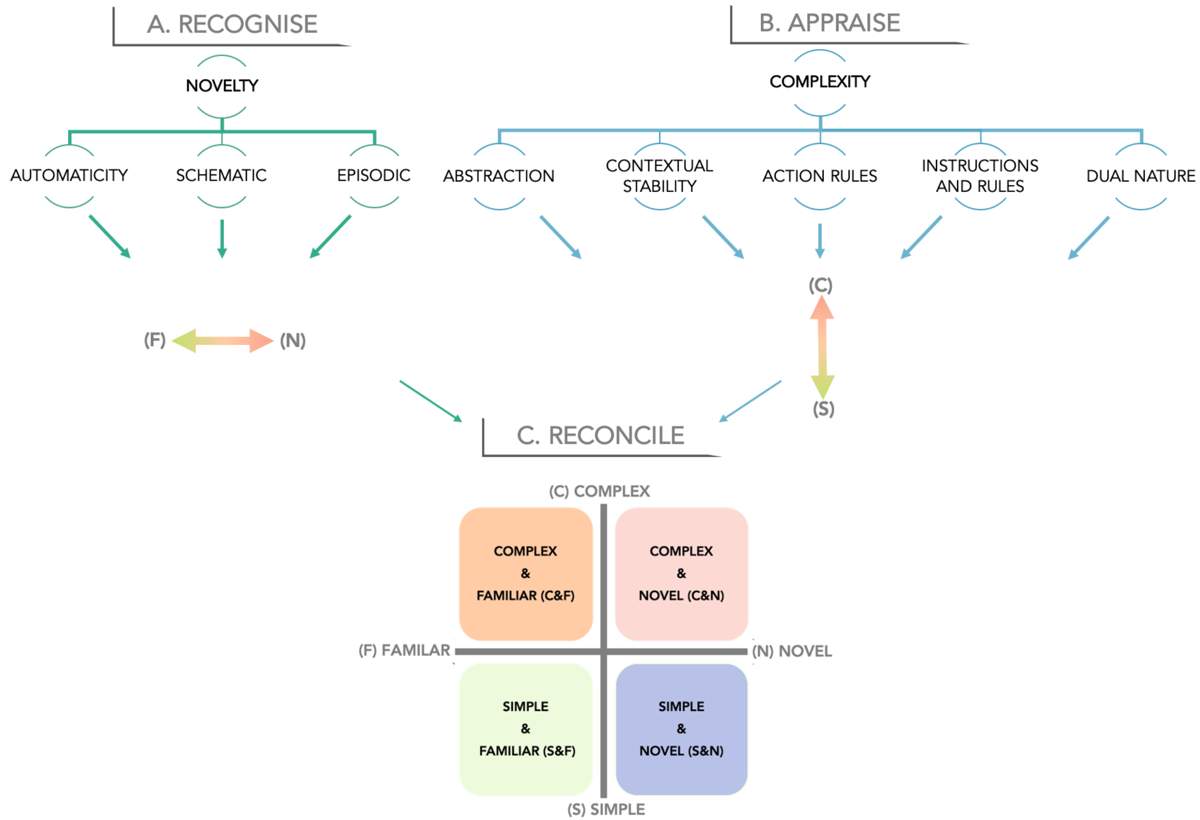
Process-driven models of higher-order cognition (Baddeley, 2007; Baddeley & Hitch, 1974; Stuss et al., 1995) provide a suitable account for the processes that must underpin the recognition of novelty within the HDM. Stuss et al. (1995) proposed five frontal supervisory processes to provide an account of familiar appraisal, and early selection of the appropriate behaviours that are required when faced with a novel situation. The support of this selection is likely afforded by WM (Baddeley, 2007), whereby the episodic buffer serves to compare current stimuli with information held in LTM stores. The Central Executive (CE) then coordinates the WM slave systems in the reconciliation of both familiar and novel information which substantiates the relative automaticity, schematic and episodic demands of the environment.

The HDM further extends the application of the process-driven models of Stuss et al. (1995) and Baddeley (2007), by proposing that their processes must be integrated with the *appraisal* of demands for complexity. In the context of process-driven models, the influence of complexity on the ability to navigate demands for novelty and familiarity of the environment has remained largely absent. While the SAS model arguably recognises the need to consider changes in the environment via the inclusion of an *'If-then…'* component, the HDM provides a detailed account of what the nature of '*if*' can be - beyond that offered by the SAS.

The HDM therefore puts forward the notion that higher-order cognition is influenced by intersecting axes of Complexity and Novelty. Each axis represents a hierarchy of environmental demand, with the horizontal axis spanning from Familiar to Novel, and the vertical axis spanning from Simple to Complex. It is important to note that the axes are conceptual in nature, and do not necessarily intersect at a point of absolute zero. As the *recognition* of task familiarity decreases, novelty in turn must increase (across the x-axis). This recognition then affects the *appraisal* in relation to task simplicity or complexity (across the y-axis). A visual representation of the HDM processes is illustrated in Figure 21.

Figure 21

The Hierarchical Demand Model (HDM)



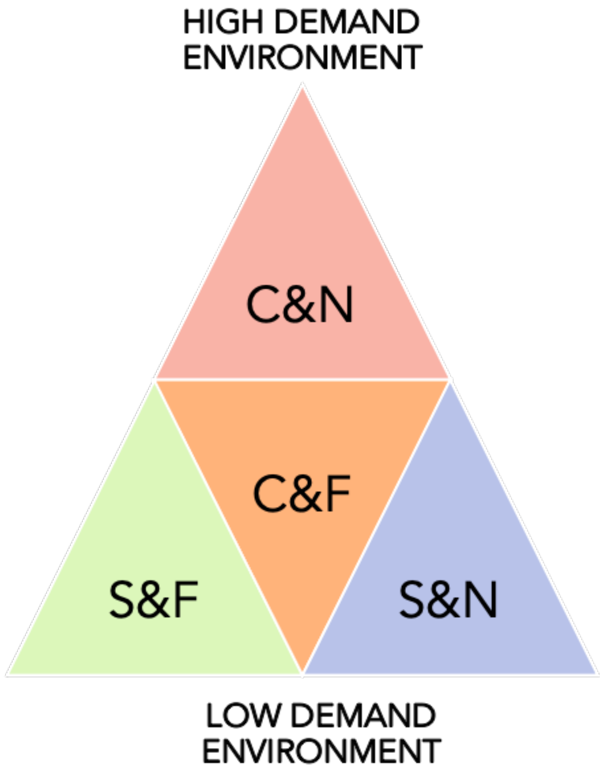
Note. The Hierarchical Demand Model (HDM) proposes that (A) in order to successfully engage in a goal-directed task the individual must *recognise* what is known, (B) *appraise* what is required, and (C) *reconcile* the difference in order to formulate an effective response. The magnitude of difference between recognition of what is known, and the appraisal of what is required, will dictate the classifications of complexity and novelty. The HDM proposes that this process is influenced by, and ultimately a product of the environment being attributable to one of four Global Demand Classifications (GDC); Simple & Familiar (S&F), Simple & Novel (S&N), Complex & Familiar (C&F), or Complex & Novel (C&N).

Once the extent of familiarity or novelty is reconciled by WM systems across the horizontal axis (coordinated by the CE), the HDM calls for the acknowledgement of the influence that complexity within the environment can then have on maintenance and retrieval. Therefore, as demand increases beyond automatic S→R relationships dictated by increases in novelty, a threshold is reached whereby behaviour must be adapted for successful performance, and the task appraisal will necessarily move towards increased complexity. In this way, the HDM moves beyond existing process-driven models which account for the recognition requirements, but do not offer the same account in regards to the appraisal of demands.

However, it is not enough to say that performance is purely dictated by recognition of familiarity or novelty affecting appraisal of task simplicity or complexity. This account of recognition and appraisal falls short if it fails to acknowledge that performance is dictated across the four quadrants of demand (S&F, S&N, C&F, C&N). The Baddeley (2007) WM model does not discriminate or provide an account of what is being retrieved from LTM storage and how this may influence performance. While the capacity of long-term memory may be infinite, and familiarity pervasive over time (Larzabal et al., 2018), recall and maintenance are limited. The HDM proposes an extension of this understanding by recognising that retrieval can be influenced by the complexity of an environment, in addition to its relative novelty. This extension is important given traditional notions that novelty ultimately bears increased demand in comparison to familiarity. The HDM instead recognises that demands for large amounts of familiar knowledge or schemas can be of equal or greater influence than novelty when required within a complex environment, as represented by S&N, C&F quadrants. These quadrants could be conceptualised as equivalent in relation to quantifiable global demand, however their functional differences add to existing theory by providing an account for why strengths and weaknesses may emerge. A conceptual illustration of the overarching continuum of demand provided by the HDM is displayed in Figure 22.

Figure 22

Conceptual Representation of the HDM



Note. This figure provides a conceptual representation of the HDM. Within the HDM, each Global Demand Classification (S&F, S&N, C&F, C&N) serves to provide a quantifiable marker for the level of demand for adaptive behaviour that is engaged when complex behaviour is undertaken. The lower levels of this pyramid reflect engagement with a low demand environment that is predominantly represented by S&F and S&N, and minor C&F features. The upper levels of this pyramid represent a high demand environment that is predominantly represented by C&N features.

While the process-driven models largely accommodate the novelty axis, the reverse holds true for domain centred models of EF. Domain centred models have arguably positioned themselves in relation to demands of complexity necessitating EF recruitment. For example, (Lezak et al., 1995) conceptualised four cognitive events and behavioural actions that are required for goal-directed behaviour in a complex environment. However, how this engagement is influenced or changed by any novelty afforded by the environment is not considered. Strict skill-based approaches to EF (e.g. Anderson et al., 2001; Miyake & Friedman, 2012; Miyake et al., 2000) have proposed that some EF skills may be considered to require the contribution of multiple skills (e.g. during Planning), or more singular skills (updating). While this approach provides an account of the skills that can be required and their respective unity or diversity, how complexity is substantiated remains elusive. Moreover, the uptake of EF Skills alone suffers similar restrictions to that of Lezak (1995), due to the lack of inclusion for how novelty functions as an influence to execution. This intersection between EF skill and demand is offered by the HDM by allowing for the identification of an environment based on the measures of these skillsets. For example, a complex task environment that is high in instruction and rules that are schematically familiar may predominantly require inhibition skills during its early stages and updating skills during later stages, due to changes in demands for novelty.

10.3 Implications

The current project sheds light on the potential over-simplification of equating neuropsychological test performance outcomes to a singular categorical EF skill by providing a quantifiable demonstration of the extent that environmental demands can account for changes to performance. This account acknowledges that multiple demand features of a task must be overcome by the recruitment and integration of cognitive, executive and control systems in order to orchestrate a suitable controlled response. While research reports that many of these tests arguably require multiple EF and cognitive systems, and that the nature of EF must inherently be adaptive, there has been a profound lack of structure available that describes how this adaptability can be applied and operationalised. Arguably, it is likely that this very challenge is what has driven robust attempts to isolate EF skill-sets in search of measurement purity, which unfortunately has only served to restrict EF to bi-modal inferences of its involvement, and has done little to progress the field. Consequently, the continued interpretation of tests under the premise that performance is constrained to a modular set of limited EF skills will only ever serve as a restricted representation of the multifaceted adaptive system that is responsive to demand.

The implications of this project are most pertinent across two key disciplines, whereby EF can begin to be positioned within its closest relative, cognitive control. In turn, cognitive control can be further explored in relation to cognitive assessment. Both approaches should adapt to embrace a synergistic approach to assessment regardless of the primary cognitive domain under investigation, as both have the capacity to influence performance. The HDM is a first attempt at integrating theories of EF and cognitive control by offering a quantification of demands that are inherent in any task. The HDM is a not a reclassification of EF, nor cognitive control, but instead an operational account that unifies the strengths of the two disciplines to account for controlled behaviour.

In addition to its theoretical and conceptual contributions, this project can offer a test battery for future clinicians to explore cognitive control and EF in a similar manner that was employed by this project. Alternatively, if the researcher or clinician does not want to employ an extensive battery, tests can be administered per Global Demand of interest. For example, if understanding is required for how capable an individual is to navigate the demands for *recognition*, *appraisal* and *reconciling* under complex demand tasks, all tasks within this project identified to encompass a complex global demand can be administered. This DCS assessment of each task across the four quadrants of the HDM embeds considerations of familiarity and novelty to account for the influence of both axes. Due to the hierarchal nature of the HDM, inherent in C&F and C&N are performance under outcomes relative to S&F and

S&N, and their influences conceptualised. If unusual performance is detected, the clinician may then progress down the hierarchy of demand to identify under which GDC any failures or disturbances to adaptation become apparent. For example, failures to recognise and integrate, and then to apply behaviour in response to the familiar demands of the task environment may account for lower performance. The researcher is then able to assess whether issues with the adaptation of familiarity mediate variations in demands for complexity. The DCS provides a useful tool to the researcher that can be applied to any given test to identify its GDC. This has the potential to extend beyond the realm of cognitive control and EF, to include other areas of cognition.

10.4 Limitations & Future Directions

It could be argued that the current study is contradictory to its critique of previous research by categorising performance from a set of neuropsychological tests. However, this argument is only applicable when one test is utilised alone and the outcome is equated to a singular performance modality. This thesis employed an evidence-based framework of demand features to assess performance across a test battery that was focused on promoting variability in performance in order to account for commonalities and differences. Moreover, a novel approach was taken by seeking to identify demands within a task to observe what this performance may represent, which was then inferred to a performance descriptor that is not singular in its definition. The GDC utilised in this study represents the summative classification of demands that ultimately vary depending on the uniqueness of the testing environment. This is in considerable contrast to the selective approach often taken by the EF literature, whereby tests are labelled as measuring a purported construct, but differ in the paradigm, administration and scoring, without providing an empirical account for how these differences may influence the performance. Conversely, studies using the same paradigms or scoring may yield different outcomes with limited explanation. This singular approach provides a limited representation of performance, whereas this study serves to represent the variability of demands that exist to challenge human ability.

Regardless of the additional performance variance that was able to be explained by the current project there remains the issue of task impurity, but added is the notion that the more complex a cognitive task becomes, the less pure the demands themselves are likely to be. Individuals approach solving complex problems from a variety of different perspectives, thereby creating a degree of “demand impurity”. However, when coupled with performance inferences from EF, this amount of impurity can be reduced and a more fulsome account of performance offered. Similarly, it could be argued that this thesis has simplified a richer ontology of cognitive control to two axes of demand. However, by use of the DCS this ontology is preserved via the demarcation of demand factors that reflect current knowledge available within the cognitive control field. As this field develops and refines our understanding of the nature of the inner mechanisms of these demand factors, refinement to their representation within the DCS must also follow.

The current project presents a novel framework that despite being evidence-based, requires further investigation and replication to discern the strength of its application and utility to the wider neuropsychological setting. The development and application of the DCS was triangulated between the research and registered practitioners. However, further validation studies are required to determine the extent of the interrater test-rest reliability and validity of classification. While the framework was grounded in neuroscientific, neuropsychological and cognitive theory, the addition of neurological imaging may have served as a beneficial adjunct to support the inferences made in regard to the framework being a representation of the rostro-caudal gradient of cognitive control. The use of this technology was outside of the scope of the current project, as the framework was first required to be developed and applied to behavioural measures to explore whether or not this methodology could identify and explain performance on these tests.

The theoretical development of the DCS was informed by studies that were by and large conducted amongst healthy adult samples, as was also the case with this project. Also, similar to many clinically focused studies, the sample size of the current study was only able to fulfil the minimum requirements for statistical analysis to enable the generalisation of the current findings to the population. The study accounted for this by adopting a statistical methodology that enabled it to increase the free parameters and reduce the risk of error. While this approach is considered robust, convention appropriately dictates that alternative analysis in the form of SEM must also be conducted to ensure the validity and replicability of the claims made here.

This project has emphasised the task environment influence and not the behavioural response tendency of the individual. It essentially measures the cognitive demands on the individual, and not the manner in which someone goes about completion. Numerous other influences on performance may be elicited by the task environment. For example, speed or accuracy is often embedded as an outcome measurement of performance, and the individual would need to consider this requirement during the reconciliation of a response. It is acknowledged that individual differences in task approach may produce a 'trade-off' of speed to improve accuracy (or vice versa), which may capture efficiency, but does not influence the overall need for adaptability. Another influence on performance may be cognitive modality, where an individual preferences verbal strategy over visual strategies. However, the influence of these is on the outcome or efficiency, and not the demand.

10.5 Conclusion and Future Directions

Human beings possess the sophisticated capacity to engage in behaviour that is both controlled and adaptive to appropriately mobilise within many different environments. Neuropsychological researchers are charged with the responsibility to establish methodologies that are able to appropriately capture and measure this capacity in the clinical setting. This project demonstrated a successful approach for quantifying the demand amongst neuropsychological testing apparatus with respect to the complexity and novelty that can exist within the testing environment. This approach served as an operationalisation of the efforts of neuropsychological, behavioural and neuroscientific research that has investigated the influence that these demands can have on controlled goal-directed behaviour.

Both the HDM and its agent, the DCS, may also serve to provide further clarity towards the selection of appropriate measures of cognitive performance. Many researchers coin the term ‘complex’ and ‘novel’ when presenting their rationale for test selection and provide little elaboration on (a) what is meant by the term complex, (B) the influence this complexity can have upon performance. While the HDM provides an account of the nature of the environment with respect to how an individual must approach different demands, this is only part of a much larger and intricate account of controlled behaviour. Promising lines of research involving the genetic determinants of complex cognition that are emerging may provide further clarity into the nature of individual differences in higher-order cognitive systems, and when the capacity to respond to specific demands can be influenced by biological determination. For example, laterality of complexity and novelty may elicit different activations within the frontal cortices. Application of the DCS to ascertain the tasks that primarily engage these activations should be adopted to support the notion that the ability to engage controlled behaviour is ultimately multiaxial, with individual factors such as age, intelligence, SES, physiological and psychological illness influencing the response that is enacted. However, the conception of the HDM provides the future researcher with the knowledge of how the testing environment can influence this response and provides a platform whereby these individual differences can be further elucidated.

References

Aben, B., Calderon, C. B., Van der Cruyssen, L., Picksak, D., Van den Bussche, E., & Verguts, T. (2019). Context-dependent modulation of cognitive control involves different temporal profiles of fronto-parietal activity. *Neuroimage, 189*, 755-762. https://doi.org/10.1016/j.neuroimage.2019.02.004

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative Analysis of Verbal Fluency Output: Review and Comparison of Several Scoring Methods. *Assessment, 8*(3), 323-338. https://doi.org/10.1177/107319110100800308

Altmann, E. M. (2004). Advance Preparation in Task Switching. *Psychological science, 15*(9), 616-622. https://doi.org/10.1111/j.0956-7976.2004.00729.x

Altmann, E. M., & Gray, W. D. (2002). Forgetting to Remember: The Functional Relationship of Decay and Interference. *Psychological science, 13*(1), 27-33. https://doi.org/10.1111/1467-9280.00405

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

Andeson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychol Bull, 103*(3), 411-423. https://doi.org/10.1037/0033-2909.103.3.411

Anderson, J. C., & Gerbing, D. W. (1992). Assumptions and Comparative Strengths of the Two-Step Approach: Comment on Fornell and Yi. *20*(3), 321-333. https://doi.org/10.1177/0049124192020003002

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

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

Anderson, P., & Reidy, N. (2012). Assessing executive function in preschoolers. *Neuropsychol Rev, 22*(4), 345-360. https://doi.org/10.1007/s11065-012-9220-3

Anderson, V., Jacobs, R., & Anderson, P. (2011). *Executive Functions and the Frontal Lobes: A Lifespan Perspective*. Taylor & Francis. http://books.google.com.au/books?id=f8p5AgAAQBAJ

Arbuckle, J. L. (2013). *IBM SPSS Amos 22 User's Guide*. Microsoft Corporation.

Asato, M. R., Sweeney, J. A., & Luna, B. (2006). Cognitive processes in the development of TOL performance. *Neuropsychologia, 44*(12), 2259-2269. https://doi.org/http://doi.org/10.1016/j.neuropsychologia.2006.05.010

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes In K. W. Spence & J. T. Spence (Eds.), *Psychology of Learning and Motivation* (Vol 2., pp 89-195). Academic Press. https://doi.org/https://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology A, 49*(1), 5-28. https://doi.org/10.1080/713755608

Baddeley, A. (1998). The central executive: a concept and some misconceptions. *J Int Neuropsychol Soc, 4*(5), 523-526. https://doi.org/10.1017/S135561779800513X

Baddeley, A. (2007). *Working Memory, Thought, and Action*. OUP Oxford.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). Academic Press. https://doi.org/https://doi.org/10.1016/S0079-7421(08)60452-1

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in cognitive sciences, 12*(5), 193-200. https://doi.org/http://dx.doi.org/10.1016/j.tics.2008.02.004

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of cognitive neuroscience, 19*(12), 2082-2099. https://doi.org/10.1162/jocn.2007.19.12.2082

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical?. *Nature Reviews Neuroscience, 10*(9), 659-669. https://doi.org/10.1038/nrn2667

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron, 66*(2), 315-326. https://doi.org/10.1016/j.neuron.2010.03.025

Badre, D., & Nee, D. E. (2018). Frontal Cortex and the Hierarchical Control of Behavior. *Trends Cogn Sci, 22*(2), 170-188. https://doi.org/10.1016/j.tics.2017.11.005

Bahlmann, J., Blumenfeld, R. S., & D'Esposito, M. (2015). The Rostro-Caudal Axis of Frontal Cortex Is Sensitive to the Domain of Stimulus Information. *25*(7), 1815-1826. https://doi.org/10.1093/cercor/bht419

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S., & Robbins, T. W. (1996). Neural systems engaged by planning: a PET study of the

Tower of London task. *Neuropsychologia, 34*(6), 515-526. https://doi.org/10.1016/0028-3932(95)00133-6

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

Barceló, F., & Cooper, P. S. (2018). Quantifying Contextual Information For Cognitive Control. *Frontiers in Psychology, 9*, 1693-1693. https://doi.org/10.3389/fpsyg.2018.01693

Barcelo, F., Escera, C., Corral, M. J., & Periáñez, J. A. (2006). Task Switching and Novelty Processing Activate a Common Neural Network for Cognitive Control. *Journal of cognitive neuroscience, 18*(10), 1734-1748. https://doi.org/10.1162/jocn.2006.18.10.1734

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues. In R. Wyer & T. Srull (Eds.), *Handbook of Social Cognition*. Lawrence Erlbaum.

Bargh, J. A., & Morsella, E. (2008). The Unconscious Mind. *Perspectives on Psychological Science, 3*(1), 73-79. https://doi.org/10.1111/j.1745-6916.2008.00064.x

Barkley, R. A. The Executive Functions and Self-Regulation: An Evolutionary Neuropsychological Perspective. *Neuropsychology Review, 11*(1), 1-29. https://doi.org/10.1023/a:1009085417776

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

Barkley, R. A. (2012). *Executive Functions: What They Are, How They Work, and Why They Evolved*. Guilford Press.

Barnard, L., Muldoon, K., Hasan, R., O'Brien, G., & Stewart, M. (2008). Profiling executive dysfunction in adults with autism and comorbid learning disability. *Autism, 12*(2), 125-141. https://doi.org/10.1177/1362361307088486

Barry, D., Bates, M. E., & Labouvie, E. (2008). FAS and CFL forms of verbal fluency differ in difficulty: a meta-analytic study. *Appl Neuropsychol, 15*(2), 97-106. https://doi.org/10.1080/09084280802083863

Bartlett, F. C., & Burt, C. (1933). Remembering: A Study In Experimental and Social Psychology. *British Journal of Educational Psychology, 3*(2), 187-192. https://doi.org/10.1111/j.2044-8279.1933.tb02913.x

Bellebaum, C., & Daum, I. (2007). Cerebellar involvement in executive control. *The Cerebellum, 6*(3), 184-192. https://doi.org/10.1080/14734220601169707

Bender, A. D., Filmer, H. L., Garner, K. G., Naughtin, C. K., & Dux, P. E. (2016). On the relationship between response selection and response inhibition: An individual differences approach [journal article]. *Attention, Perception, & Psychophysics*, 1-13. https://doi.org/10.3758/s13414-016-1158-8

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

Berg, W. K., Byrd, D. L., McNamara, J. P., & Case, K. (2010). Deconstructing the tower: parameters and predictors of problem difficulty on the Tower of London task. *Brain Cogn, 72*(3), 472-482. https://doi.org/10.1016/j.bandc.2010.01.002

Berkman, E. T. (2018). The Neuroscience of Goals and Behavior Change. *Consulting psychology journal, 70*(1), 28-44. https://doi.org/10.1037/cpb0000094

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev, 81*(6), 1641-1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bhandari, A., & Badre, D. (2018). Learning and transfer of working memory gating policies. *Cognition, 172*, 89-100. https://doi.org/10.1016/j.cognition.2017.12.001

Bisho, 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 & Psychiatry & Allied Disciplines, 42*(4), 551. https://doi.org/10.1111/1469-7610.00749

Bocanegra, B. R., & Hommel, B. (2014). When Cognitive Control Is Not Adaptive. *25*(6), 1249-1255. https://doi.org/10.1177/0956797614528522

Bora, E., Hıdıroğlu, C., Özerdem, A., Kaçar, Ö. F., Sarısoy, G., Civil Arslan, F., Aydemir, Ö., Cubukcuoglu Tas, Z., Vahip, S., Atalay, A., Atasoy, N., Ateşci, F., & Tümkaya, S. (2016). Executive dysfunction and cognitive subgroups in a large sample of euthymic patients with bipolar disorder. *European Neuropsychopharmacology*. https://doi.org/http://dx.doi.org/10.1016/j.euroneuro.2016.04.002

Bott, N. T., Johnson, E. T., Schuff, N., Galifianakis, N., Subas, T., Pollock, J., Pressman, P., Kramer, J. H., & Possin, K. L. (2014). Sensitive measures of executive dysfunction in non-demented Parkinson's disease. *Parkinsonism & Related Disorders, 20*(12), 1430-1433. https://doi.org/http://dx.doi.org/10.1016/j.parkreldis.2014.10.007

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in cognitive sciences, 12*(5), 201-208. https://doi.org/http://dx.doi.org/10.1016/j.tics.2008.02.009

Bowden, S. C., & Smith, L. C. (1994). What does the Austin Maze measure? *Australian Psychologist, 29*(1), 34-37. http://0-search.ebscohost.com.library.vu.edu.au/login.aspx?direct=true&db=psyh&AN=1994-32016-001&site=ehost-live

Brass, M. (2002). The Role of the Frontal Cortex in Task Preparation. *Cerebral Cortex, 12*(9), 908-914. https://doi.org/10.1093/cercor/12.9.908

Brass, M., Liefooghe, B., Braem, S., & De Houwer, J. (2017). Following new task instructions: Evidence for a dissociation between knowing and doing. *Neurosci Biobehav Rev, 81*, 16-28. https://doi.org/10.1016/j.neubiorev.2017.02.012

Brickenkamp, R., & Zillmer, E. (1998). *The d2 test of Attention* (1st edition ed.). Hogrefe Publishing.

Brown, L. A., Brockmole, J. R., Gow, A. J., & Deary, I. J. (2012). Processing Speed and Visuospatial Executive Function Predict Visual Working Memory Ability in Older Adults. *Experimental Aging Research, 38*(1), 1-19. https://doi.org/10.1080/0361073X.2012.636722

Brown, S. D., Marley, A. A., Donkin, C., & Heathcote, A. (2008). An integrated model of choices and response times in absolute identification. *Psychol Rev, 115*(2), 396-425. https://doi.org/10.1037/0033-295x.115.2.396

Bub, D. N., Masson, M. E. J., & Lalonde, C. E. (2006). Cognitive Control in Children. *Psychological science, 17*(4), 351-357. https://doi.org/10.1111/j.1467-9280.2006.01710.x

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *J Child Psychol Psychiatry, 45*(4), 743-754. https://doi.org/10.1111/j.1469-7610.2004.00268.x

Bunge, S. A. (2000). A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences, 97*(7), 3573-3578. https://doi.org/10.1073/pnas.97.7.3573

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

Burgess, P., Alderman, N., Forbes, C., Costello, A., Coates, L., Dawson, D., Anderson, N., Gilbert, S., Dumontheil, I., & Channon, S (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. Journal of the International Neuropsychological Society, 12(2), 194-209. https://doi.org/10.1017/S1355617706060310

Burke, L. E., Swigart, V., Warziski Turk, M., Derro, N., & Ewing, L. J. (2009). Experiences of self-monitoring: successes and struggles during treatment for weight loss. *Qualitative health research, 19*(6), 815-828. https://doi.org/10.1177/1049732309335395

Byrne, B. M. (2016). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming, Third Edition*. Taylor & Francis. https://books.google.com.au/books?id=d81TDAAAQBAJ

Cangur, S., & Ercan, I. (2015). Comparison of Model Fit Indices Used in Structural Equation Modeling Under Multivariate Normality. *Journal of Modern Applied Statistical Methods, 14*(1), 152-167.

Cardillo, R., Mammarella, I. C., Garcia, R. B., & Cornoldi, C. (2017). Local and global processing in block design tasks in children with dyslexia or nonverbal learning disability. *Res Dev Disabil, 64*, 96-107. https://doi.org/10.1016/j.ridd.2017.03.011

Carver, C. S., & Scheier, M. F. (2001). *On the Self-Regulation of Behavior*. Cambridge University Press.

Chan, A. S., Cheung, M.-c., Han, Y. M. Y., Sze, S. L., Leung, W. W., Man, H. S., & To, C. Y. (2009). Executive function deficits and neural discordance in children with Autism Spectrum Disorders. *Clinical Neurophysiology, 120*(6), 1107-1115. https://doi.org/http://dx.doi.org/10.1016/j.clinph.2009.04.002

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. https://doi.org/http://dx.doi.org/10.1016/j.acn.2007.08.010

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Arch Clin Neuropsychol, 21*(3), 217-227. https://doi.org/10.1016/j.acn.2005.12.002

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Chen, G., Adleman, N. E., Saad, Z. S., Leibenluft, E., & Cox, R. W. (2014). Applications of multivariate modeling to neuroimaging group analysis: A comprehensive alternative to univariate general linear model. *Neuroimage, 99*, 571-588. https://doi.org/http://dx.doi.org/10.1016/j.neuroimage.2014.06.027

Christensen, W., Sutton, J., & Bicknell, K. (2019). Memory systems and the control of skilled action. *Philosophical Psychology, 32*(5), 692-718. https://doi.org/10.1080/09515089.2019.1607279

Coakes, S. J. (2012). *SPSS version 20.0 for Windows: Analysis without anguish*. Wiley.

Cohen-Kdoshay, O., & Meiran, N. (2007). The representation of instructions in working memory leads to autonomous response activation: evidence from the first trials in the flanker paradigm. *Q J Exp Psychol, 60*(8), 1140-1154. https://doi.org/10.1080/17470210600896674

Cohen-Kdoshay, O., & Meiran, N. (2009). The Representation of Instructions Operates Like a Prepared Reflex. *56*(2), 128-133. https://doi.org/10.1027/1618-3169.56.2.128

Cole, M., Etzel, J., Zacks, J., Schneider, W., & Braver, T. (2011). Rapid Transfer of Abstract Rules to Novel Contexts in Human Lateral Prefrontal Cortex. *Frontiers in Human Neuroscience, 5*(142). https://doi.org/10.3389/fnhum.2011.00142

Cole, M. W., Laurent, P., & Stocco, A. (2013). Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 13*(1), 1-22. https://doi.org/10.3758/s13415-012-0125-7

Cole, M. W., Patrick, L. M., Meiran, N., & Braver, T. S. (2018). A role for proactive control in rapid instructed task learning. *Acta Psychol (Amst), 184*, 20-30. https://doi.org/10.1016/j.actpsy.2017.06.004

Collette, F., Hogge, M., Salmon, E., & Van der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience, 139*(1), 209-221. https://doi.org/10.1016/j.neuroscience.2005.05.035

Collette, F., Olivier, L., Van der Linden, M., Laureys, S., Delfiore, G., Luxen, A., & Salmon, E. (2005). Involvement of both prefrontal and inferior parietal cortex in dual-task performance. *Brain Res Cogn Brain Res, 24*(2), 237-251. https://doi.org/10.1016/j.cogbrainres.2005.01.023

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. *Hum Brain Mapp, 25*(4), 409-423. https://doi.org/10.1002/hbm.20118

Collins, A. G., & Frank, M. J. (2013). Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol Rev, 120*(1), 190-229. https://doi.org/10.1037/a0030852

Connolly, J., McNulty, J. P., Boran, L., Roche, R. A., Delany, D., & Bokde, A. L. (2016). Identification of Resting State Networks Involved in Executive Function. *Brain Connect, 6*(5), 365-374. https://doi.org/10.1089/brain.2015.0399

Cooper, R., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cogn Neuropsychol, 17*(4), 297-338. https://doi.org/10.1080/026432900380427

Cothran, D. L., & Larsen, R. (2008). Comparison of inhibition in two timed reaction tasks: the color and emotion Stroop tasks. *The Journal of psychology, 142*(4), 373-385. https://doi.org/10.3200/JRLP.142.4.373-385

Coutlee, C. G., & Huettel, S. A. (2012). The functional neuroanatomy of decision making: prefrontal control of thought and action. *Brain research, 1428*, 3-12. https://doi.org/10.1016/j.brainres.2011.05.053

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684. https://doi.org/10.1016/s0022-5371(72)80001-x

Crowe, S. F., Barclay, L., Brennan, S., Farkas, L., Gould, E., Katchmarsky, S., & Vayda, S. (1999). The cognitive determinants of performance on the Austin Maze. *Journal of the International Neuropsychological Society, 5*(1), 1-9. https://doi.org/10.1017/s1355617799511016

Dagher, A., Owen, A. M., Boecker, H., & Brooks, D. J. (1999). Mapping the network for planning: a correlational PET activation study with the Tower of London task. *Brain, 122*, 1973-1987. https://doi.org/10.1093/brain/122.10.1973

Darby, D. (2014a). *Span Tests (Version 1.2)([iPad Application]*. WhiteAnt Occasional Publishing. https://apps.apple.com/bg/app/spantests/id479165359

Darby, D. (2014b). *Austin Maze (Version 2.3) [iPad Application]*. WhiteAnt Occasional Publishing. https://apps.apple.com/bg/app/austinmaze/id492176361

Darby, D., & Walsh, K. W. (2005). *Walsh's Neuropsychology: A Clinical Approach*. Elsevier Churchill Livingstone.

Davis, T., Goldwater, M., & Giron, J. (2016). From Concrete Examples to Abstract Relations: The Rostrolateral Prefrontal Cortex Integrates Novel Examples into Relational Categories. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhw099

Dawson, P., & Guare, R. (2010). *Executive Skills in Children and Adolescents, Second Edition: A Practical Guide to Assessment and Intervention*. Guilford Publications.

DeFelipe, J. (2011). The Evolution of the Brain, the Human Nature of Cortical Circuits, and Intellectual Creativity. *Frontiers in Neuroanatomy, 5*(29). https://doi.org/10.3389/fnana.2011.00029

Desrochers, T. M., Collins, A. G. E., & Badre, D. (2019). Sequential Control Underlies Robust Ramping Dynamics in the Rostrolateral Prefrontal Cortex. *The Journal of Neuroscience, 39*(8), 1471. https://doi.org/10.1523/JNEUROSCI.1060-18.2018

Dezfouli, A., & Balleine, B. W. (2013). Actions, Action Sequences and Habits: Evidence That Goal-Directed and Habitual Action Control Are Hierarchically Organized. *PLoS Computational Biology, 9*(12), e1003364. https://doi.org/10.1371/journal.pcbi.1003364

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology, 64*(1), 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dickinson, A. (1985). Actions and Habits: The Development of Behavioural Autonomy. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 308*(1135), 67-78. www.jstor.org/stable/2396284

Dickstein, S. G., Bannon, K., Xavier Castellanos, F., & Milham, M. P. (2006). The neural correlates of attention deficit hyperactivity disorder: an ALE meta-analysis [Article]. *Journal of Child Psychology & Psychiatry, 47*(10), 1051-1062. https://doi.org/10.1111/j.1469-7610.2006.01671.x

Ding, L., Velicer, W. F., & Harlow, L. L. (1995). Effects of estimation methods, number of indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling: A Multidisciplinary Journal, 2*(2), 119-143. https://doi.org/10.1080/10705519509540000

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

Domenech, P., & Koechlin, E. (2015). Executive control and decision-making in the prefrontal cortex. *Current Opinion in Behavioral Sciences, 1*, 101-106. https://doi.org/http://dx.doi.org/10.1016/j.cobeha.2014.10.007

Donnarumma, F., Maisto, D., & Pezzulo, G. (2016). Problem Solving as Probabilistic Inference with Subgoaling: Explaining Human Successes and Pitfalls in the Tower of Hanoi [Article]. *PLoS Computational Biology, 12*(4), 1-30. https://doi.org/10.1371/journal.pcbi.1004864

Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron, 80*(1), 35-50. https://doi.org/10.1016/j.neuron.2013.09.015

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn Psychol, 30*(3), 257-303. https://doi.org/10.1006/cogp.1996.0008

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. https://doi.org/http://dx.doi.org/10.1016/S0166-2236(00)01633-7

Durston, S., Tottenham, N. T., Thomas, K. M., Davidson, M. C., Eigsti, I.-M., Yang, Y., Ulug, A. M., & Casey, B. J. (2003). Differential patterns of striatal activation in young children with and without ADHD. *Biological Psychiatry, 53*(10), 871-878. https://doi.org/10.1016/S0006-3223(02)01904-2

Elliott, R. (2003). Executive functions and their disorders: Imaging in clinical neuroscience. *British Medical Bulletin, 65*(1), 49-59. https://doi.org/10.1093/bmb/65.1.49

Escera, C., Alho, K., Winkler, I., & Naatanen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *J Cogn Neurosci, 10*(5), 590-604. https://doi.org/10.1162/089892998562997

Farooqui, A. A., & Manly, T. (2019). We do as we construe: extended behavior construed as one task is executed as one cognitive entity. *Psychological Research, 83*(1), 84-103. https://doi.org/10.1007/s00426-018-1051-2

Farooqui, A. A., Mitchell, D., Thompson, R., & Duncan, J. (2012). Hierarchical Organization of Cognition Reflected in Distributed Frontoparietal Activity. *Journal Of Neuroscience, 32*(48), 17373-17381. https://doi.org/10.1523/JNEUROSCI.0598-12.2012

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, 110*(41), 16616-16621. https://doi.org/10.1073/pnas.1315235110

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

Fetsch, C. R. (2016). The importance of task design and behavioral control for understanding the neural basis of cognitive functions. *Current Opinion in Neurobiology, 37*, 16-22. https://doi.org/10.1016/j.conb.2015.12.002

Field, A. (2007). *Discovering Statistics Using SPSS*. SAGE Publications.

Fischer, R., & Plessow, F. (2015). Efficient multitasking: parallel versus serial processing of multiple tasks [Review]. *Frontiers in Psychology, 6*(1366). https://doi.org/10.3389/fpsyg.2015.01366

Fishbach, A., Dhar, R., & Zhang, Y. (2006). Subgoals as substitutes or complements: the role of goal accessibility. *J Pers Soc Psychol, 91*(2), 232-242. https://doi.org/10.1037/0022-3514.91.2.232

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: updating, inhibition, shifting, and access. *J Clin Exp Neuropsychol, 26*(7), 874-890. https://doi.org/10.1080/13803390490510680

Fournier-Vicente, S., Larigauderie, P., & Gaonac'h, D. (2008). More dissociations and interactions within central executive functioning: a comprehensive latent-variable analysis. *Acta Psychol (Amst), 129*(1), 32-48. https://doi.org/10.1016/j.actpsy.2008.04.004

Friedman, N. P., & Miyake, A. (2016). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*. https://doi.org/http://dx.doi.org/10.1016/j.cortex.2016.04.023

Fuster, J. M. (2000a). Executive frontal functions. *Exp Brain Res,* 133, 66–70. https://doi.org/10.1007/s002210000401

Fuster, J. M. (2000b). Prefrontal neurons in networks of executive memory. *Brain Res Bull, 52*(5), 331-336. https://doi.org/10.1016/S0361-9230(99)00258-0

Fuster, J. M. (2001). The Prefrontal Cortex—An Update: Time Is of the Essence. *Neuron, 30*(2), 319-333. https://doi.org/http://dx.doi.org/10.1016/S0896-6273(01)00285-9

Fuster, J. M. (2002). Frontal lobe and cognitive development. *J Neurocytol, 31*(3-5), 373-385.

Fuster, J. M. (2017). Chapter 8 - Prefrontal Cortex in Decision-Making: The Perception–Action Cycle. In J. C. Dreher & L. Tremblay (Eds.), *Decision Neuroscience* (pp. 95-105). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-805308-9.00008-7

García-Madruga, J. A., Gómez-Veiga, I., & Vila, J. Ó. (2016). Executive Functions and the Improvement of Thinking Abilities: The Intervention in Reading Comprehension. *Frontiers in Psychology, 7*(58). https://doi.org/10.3389/fpsyg.2016.00058

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia, 53*, 104-114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Gilbert, S. J., Bird, G., Brindley, R., Frith, C. D., & Burgess, P. W. (2008). Atypical recruitment of medial prefrontal cortex in autism spectrum disorders: An fMRI study of two executive function tasks. *Neuropsychologia, 46*, 2281-2291. https://doi.org/10.1016/j.neuropsychologia.2008.03.025

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

Goebel, S., Fischer, R., Ferstl, R., & Mehdorn, H. M. (2009). Normative data and psychometric properties for qualitative and quantitative scoring criteria of the Five-point Test. *Clin Neuropsychol, 23*(4), 675-690. https://doi.org/10.1080/13854040802389185

Goldstein, S., & Naglieri, J. A. (2013). *Handbook of Executive Functioning*. Springer New York.

Gross, R. G., & Grossman, M. (2010). Executive Resources. *Continuum: Lifelong Learning in Neurology, 16*, 140-152. https://doi.org/10.1212/01.con.0000368266.46038.0e

Groth-Marnat, G. (2003). *Handbook of Psychological Assessment*. Wiley.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate Data Analysis*. Pearson Education Limited.

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.* Scientific Software International, Inc

Hanus, Z., Tatasinska, J., & Osypiuk, Z. On the small sample properties of variants of Mardia's and Srivastava's kurtosis-based tests for multivariate normality. *Biometrical Letters, 49*(2), 159-175. https://doi.org/10.2478/bile-2013-0012

Hartstra, E., Waszak, F., & Brass, M. (2012). The implementation of verbal instructions: Dissociating motor preparation from the formation of stimulus–response associations. *63*(3), 1143-1153. https://doi.org/10.1016/j.neuroimage.2012.08.003

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: evidence for set-level control. *J Exp Psychol Hum Percept Perform, 37*(6), 1898-1914. https://doi.org/10.1037/a0024662

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in cognitive sciences, 8*(1), 26-32. https://doi.org/http://dx.doi.org/10.1016/j.tics.2003.11.003

Höchli, B., Brügger, A., & Messner, C. (2018). How Focusing on Superordinate Goals Motivates Broad, Long-Term Goal Pursuit: A Theoretical Perspective. *Frontiers in Psychology, 9*, 1879-1879. https://doi.org/10.3389/fpsyg.2018.01879

Hommel, B. (1998). Automatic stimulus–response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 24*(5), 1368-1384. https://doi.org/10.1037/0096-1523.24.5.1368

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science, 310*(5754), 1680-1683. https://doi.org/10.1126/science.1115327

Hu, L. T., & 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. https://doi.org/10.1080/10705519909540118

Huey, E. D., Goveia, E. N., Paviol, S., Pardini, M., Krueger, F., Zamboni, G., Tierney, M. C., Wassermann, E. M., & Grafman, J. (2009). Executive dysfunction in frontotemporal dementia and corticobasal syndrome. *Neurology, 72*(5), 453-459. https://doi.org/10.1212/01.wnl.0000341781.39164.26

Hugdahl, K., Raichle, M. E., Mitra, A., & Specht, K. (2015). On the existence of a generalized non-specific task-dependent network [Hypothesis & Theory]. *Frontiers in Human Neuroscience, 9*. https://doi.org/10.3389/fnhum.2015.00430

Hughes, C., & Graham, A. (2002). Measuring Executive Functions in Childhood: Problems and Solutions? *Child and Adolescent Mental Health, 7*(3), 131-142. https://doi.org/10.1111/1475-3588.00024

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017-2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

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. https://doi.org/10.1177/107319119700400305

Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics [Article]. *Journal of Consumer Psychology, 20*, 90-98. https://doi.org/10.1016/j.jcps.2009.09.003

Jaeggi, S. M., Seewer, R., Nirkko, A. C., Eckstein, D., Schroth, G., Groner, R., & Gutbrod, K. (2003). Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: functional magnetic resonance imaging study. *Neuroimage, 19*(2), 210-225. https://doi.org/http://dx.doi.org/10.1016/S1053-8119(03)00098-3

Janowich, J. R., & Cavanagh, J. F. (2019). Immediate versus delayed control demands elicit distinct mechanisms for instantiating proactive control. *Cognitive, affective & behavioral neuroscience, 19*(4), 910-926. https://doi.org/10.3758/s13415-018-00684-x

Jeon, H.A. (2014). Hierarchical processing in the prefrontal cortex in a variety of cognitive domains. *Frontiers in Systems Neuroscience, 8*(223). https://doi.org/10.3389/fnsys.2014.00223

Jeon, H.A., & Friederici, A. D. (2015). Degree of automaticity and the prefrontal cortex. *Trends in cognitive sciences, 19*(5), 244-250. https://doi.org/10.1016/j.tics.2015.03.003

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: a guide to the program and applications*. SPSS Inc.

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

Kaller, C. P., Rahm, B., Köstering, L., & Unterrainer, J. M. (2011). Reviewing the impact of problem structure on planning: A software tool for analyzing tower tasks. *Behavioural Brain Research, 216*(1), 1-8. http://doi.org/10.1016/j.bbr.2010.07.029

Kaller, C. P., Rahm, B., Spreer, J., Mader, I., & Unterrainer, J. M. (2008). Thinking around the corner: The development of planning abilities. *Brain and Cognition, 67*(3), 360-370. https://doi.org/10.1016/j.bandc.2008.02.003

Kaller, C. P., Unterrainer, J. M., Rahm, B., & Halsband, U. (2004). The impact of problem structure on planning: insights from the Tower of London task. *Cognitive Brain Research, 20*(3), 462-472. https://doi.org/10.1016/j.cogbrainres.2004.04.002

Kaller, C. P., Weiller, C., Rahm, B., Unterrainer, J. M., & Spreer, J. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cerebral Cortex, 21*(2), 307-317. https://doi.org/10.1093/cercor/bhq096

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 Exp Psychol Gen, 132*(1), 47-70. https://doi.org/10.1037/0096-3445.132.1.47

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. *Psychol Bull, 144*(11), 1147-1185. https://doi.org/10.1037/bul0000160

Kessels, R. P., van Zandvoort, M. J., Postma, A., Kappelle, L. J., & de Haan, E. H. (2000). The Corsi Block-Tapping Task: standardization and normative data. *Appl Neuropsychol, 7*(4), 252-258. https://doi.org/10.1207/s15324826an0704_8

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

Koechlin, E. (2016). Prefrontal executive function and adaptive behavior in complex environments. *Current Opinion in Neurobiology, 37*, 1-6. https://doi.org/http://dx.doi.org/10.1016/j.conb.2015.11.004

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science, 302*(5648), 1181-1185. http://www.jstor.org/stable/3835489

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in cognitive sciences, 11*(6), 229-235. https://doi.org/http://dx.doi.org/10.1016/j.tics.2007.04.005

Kohs, S. C. (1923). *Intelligence measurement : a psychological and statistical study based upon the block-design tests*. New York : The Macmillan Company. https://doi.org/https://doi.org/10.1037/11201-000

Konishi, S., Kawazu, M., Uchida, I., Kikyo, H., Asakura, I., & Miyashita, Y. (1999). Contribution of Working Memory to Transient Activation in Human Inferior Prefrontal Cortex during Performance of the Wisconsin Card Sorting Test. *Cerebral Cortex, 9*(7), 745-753. https://doi.org/10.1093/cercor/9.7.745

Kortte, K. 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. https://doi.org/10.1207/s15324826an0902_5

Kray, J., Eppinger, B., & Mecklinger, A. (2005). Age differences in attentional control: An event-related potential approach. *Psychophysiology, 42*(4), 407-416. https://doi.org/doi:10.1111/j.1469-8986.2005.00298.x

Kroger, J. K. (2002). Recruitment of Anterior Dorsolateral Prefrontal Cortex in Human Reasoning: a Parametric Study of Relational Complexity. *Cerebral Cortex, 12*(5), 477-485. https://doi.org/10.1093/cercor/12.5.477

Larzabal, C., Tramoni, E., Muratot, S., Thorpe, S. J., & Barbeau, E. J. (2018). Extremely long-term memory and familiarity after 12 years. *Cognition, 170*, 254-262. https://doi.org/10.1016/j.cognition.2017.10.009

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *J Exp Psychol Gen, 133*(3), 339-354. https://doi.org/10.1037/0096-3445.133.3.339

Levine, S. Z., Petrides, K. V., Davis, S., Jackson, C. J., & Howell, P. (2005). The Use of Structural Equation Modeling in Stuttering Research: Concepts and Directions. *Stammering research, 1*(4), 344-363. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2493410/

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological Assessment*. Oxford University Press.

Lezak, M. D., Lezak, P. E. N. P. N. M. D., & Press, O. U. (1995). *Neuropsychological Assessment*. Oxford University Press.

Liao, H.-I., Yeh, S.-L., & Shimojo, S. (2011). Novelty vs. Familiarity Principles in Preference Decisions: Task-Context of Past Experience Matters. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00043

Lie, C.-H., Specht, K., Marshall, J. C., & Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage, 30*(3), 1038-1049. https://doi.org/http://dx.doi.org/10.1016/j.neuroimage.2005.10.031

Liefooghe, B., & Verbruggen, F. (2019). On the Assimilation of Instructions: Stimulus-response Associations are Implemented but not Stimulus-task Associations. *Journal of cognition, 2*(1), 20-20. https://doi.org/10.5334/joc.78

Liefooghe, B., Wenke, D., & De Houwer, J. (2012). Instruction-based task-rule congruency effects. *J Exp Psychol Learn Mem Cogn, 38*(5), 1325-1335. https://doi.org/10.1037/a0028148

Logie, R. H. (2016). Retiring the central executive. *The Quarterly Journal of Experimental Psychology, 69*(10), 2093-2109. https://doi.org/10.1080/17470218.2015.1136657

Longman, C. S., Liefooghe, B., & Verbruggen, F. (2019). How Does the (Re)Presentation of Instructions Influence Their Implementation? *Journal of cognition, 2*(1). https://doi.org/10.5334/joc.63

Lovstad, M., Funderud, I., Endestad, T., Due-Tonnessen, P., Meling, T. R., Lindgren, M., Knight, R. T., & Solbakk, A. K. (2012). Executive functions after orbital or lateral prefrontal lesions: neuropsychological profiles and self-reported executive functions in everyday living. *Brain Inj, 26*(13-14), 1586-1598. https://doi.org/10.3109/02699052.2012.698787

Luria, A. R. (1976). *The Working Brain: An Introduction To Neuropsychology*. Basic Books.

MacCallum, R. C., Widaman, K., F. , & Zhand, S., H. S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99. https://doi.org/1082-989X/99/S3.00

Mäki-Marttunen, V., Hagen, T., & Espeseth, T. (2019). Proactive and reactive modes of cognitive control can operate independently and simultaneously. *Acta Psychol (Amst), 199*. 102891. https://doi.org/10.1016/j.actpsy.2019.102891

Mäki-Marttunen, V., Hagen, T., & Espeseth, T. (2019). Task context load induces reactive cognitive control: An fMRI study on cortical and brain stem activity. *Cognitive, Affective, & Behavioral Neuroscience, 19*(4), 945-965. https://doi.org/10.3758/s13415-019-00691-6

Manoach, D. S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B. M., Benfield, A., Edelman, R. R., & Warach, S. (1997). Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport, 8*(2), 545-549. http://www.nmr.mgh.harvard.edu/manoachlab/pubs/manoach_neurorep_97.pdf

Mazzone, M. (2015). Constructing the context through goals and schemata: top-down processes in comprehension and beyond. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00651

Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2015). The power of instructions: Proactive configuration of stimulus-response translation. *J Exp Psychol Learn Mem Cogn, 41*(3), 768-786. https://doi.org/10.1037/xlm0000063

Messer, D., Bernardi, M., Botting, N., Hill, E. L., Nash, G., Leonard, H. C., & Henry, L. A. (2018). An Exploration of the Factor Structure of Executive Functioning in Children. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.01179

Miches, L., Bucher, K., Luchinger, R., Klaver, P., Martin, E., Jeanmonod, D., & Brandeis, D. (2010). Simultaneous EEG-fMRI during a Working Memory Task: Modulations in Low and High Frequency Bands. *PLoS ONE, 5*(4). https://doi.org/10.1371/journal.pone.0010298

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci, 24*, 167-202. https://doi.org/10.1146/annurev.neuro.24.1.167

Minamoto, T., Yaoi, K., Osaka, M., & Osaka, N. (2015). The rostral prefrontal cortex underlies individual differences in working memory capacity: An approach from the hierarchical model of the cognitive control. *Cortex, 71*, 277-290. https://doi.org/http://dx.doi.org/10.1016/j.cortex.2015.07.025

Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives Of General Psychiatry, 66*(8), 811-822. https://doi.org/10.1001/archgenpsychiatry.2009.91

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Curr Dir Psychol Sci, 21*(1), 8-14. https://doi.org/10.1177/0963721411429458

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. *Cogn Psychol, 41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Modave, F., & Shokar, N. K. (2013). Non-Additive Measures: A Theoretical Approach to Medical Decision-Making. *Journal of Information Technology & Software Engineering 3*(3), 1-6. https://doi.org/10.4172/2165-7866.1000124

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychol Bull, 132*(2), 297-326. https://doi.org/10.1037/0033-2909.132.2.297

Muhle-Karbe, P. S., Duncan, J., De Baene, W., Mitchell, D. J., & Brass, M. (2016). Neural Coding for Instruction-Based Task Sets in Human Frontoparietal and Visual Cortex. *Cerebral Cortex*, bhw032. https://doi.org/10.1093/cercor/bhw032

Munck, I., M, E. (1979). *Model building in comparative education: Applications of the LISREL method to cross-national survey data*. Almqvist and Wiksell.

Murphy, C. P., Jackson, R. C., & Williams, A. M. (2018). The role of contextual information during skilled anticipation. *Q J Exp Psychol (Hove), 71*(10), 2070-2087. https://doi.org/10.1177/1747021817739201

Nee, D. E., & D'Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *Elife, 5*. https://doi.org/10.7554/eLife.12112

Nee, D. E., Jahn, A., & Brown, J. W. (2014). Prefrontal Cortex Organization: Dissociating Effects of Temporal Abstraction, Relational Abstraction, and Integration with fMRI. *Cerebral Cortex, 24*(9), 2377-2387. https://doi.org/10.1093/cercor/bht091

Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In F. N. Dempster, C. J. Brainerd, & C. J. Brainerd (Eds.), *Interference and Inhibition in Cognition* (pp. 207-261). Academic Press. https://doi.org/https://doi.org/10.1016/B978-012208930-5/50008-8

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 241-268. https://doi.org/10.3758/s13415-011-0083-5

Niki, C., Kumada, T., Maruyama, T., Tamura, M., & Muragaki, Y. (2019). Role of Frontal Functions in Executing Routine Sequential Tasks. *Frontiers in Psychology, 10*(169). https://doi.org/10.3389/fpsyg.2019.00169

Nimmo-Smith, I., Robertson, I. H., Ward, T., & Ridgeway, V. (1994). *Test of Everyday Attention (TEA)*. Pearson Clinical Assessment

Norman, D., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R. Davidson, R. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-Regulation: Advances in Research and Theory IV*. Plenum Press.

Nyongesa, M. K., Ssewanyana, D., Mutua, A. M., Chongwo, E., Scerif, G., Newton, C. R. J. C., & Abubakar, A. (2019). Assessing Executive Function in Adolescence: A Scoping Review of Existing Measures and Their Psychometric Robustness. *Frontiers in Psychology, 10*(311). https://doi.org/10.3389/fpsyg.2019.00311

Olszanowski, M., & Szostak, N. (2019). Adjustment of cognitive control to the frequency of dual-task interference. *Journal of Cognitive Psychology, 31*(5-6), 582-593. https://doi.org/10.1080/20445911.2019.1635605

Oosterlaan, J., Scheres, A., & Sergeant, J. A. (2005). Which Executive Functioning Deficits Are Associated With AD/HD, ODD/CD and Comorbid AD/HD+ODD/CD? *Journal of Abnormal Child Psychology, 33*(1), 69-85. https://doi.org/10.1007/s10802-005-0935-y

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*(4), 351-371. https://doi.org/10.1007/BF02213420

Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *J Clin Exp Neuropsychol, 33*(4), 456-470. https://doi.org/10.1080/13803395.2010.533157

Palenciano, A. F., González-García, C., Arco, J. E., Pessoa, L., & Ruz, M. (2019). Representational organization of novel task sets during proactive encoding. *bioRxiv*, 710244. https://doi.org/10.1101/710244

Palmer, T. E. (1975). The effects of contextual scenes on the identification of objects. *Mem Cognit, 3*(5), 519-526. https://doi.org/10.3758/bf03197524

Pashler, H., & Johnston, J. C. (1989). Chronometric Evidence for Central Postponement in Temporally Overlapping Tasks. *The Quarterly Journal of Experimental Psychology Section A, 41*(1), 19-45. https://doi.org/10.1080/14640748908402351

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., & Gore, J. C. (1999). An fMRI study of stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. *45*(10), 1237-1258. https://doi.org/10.1016/s0006-3223(99)00056-6

Pezzulo, G., Verschure, P. F. M. J., Balkenius, C., & Pennartz, C. M. A. (2014). The principles of goal-directed decision-making: from neural mechanisms to computation and robotics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1655), 20130470. https://doi.org/10.1098/rstb.2013.0470

Pietrzak, R., Maruff, P., Mayes, L., Roman, S., Sosa, J., & Snyder, P. (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. https://doi.org/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. https://doi.org/10.1080/00207450390162164

Pisula, W., Modlinska, K., & Chrzanowska, A. (2019). Behavioural Response to the Environmental Changes of Various Types in Lister-Hooded Male Rats. *Scientific Reports, 9*(1), 7111. https://doi.org/10.1038/s41598-019-42924-1

Posner, M. I., & Snyder, C. R. R. (2004). Attention and Cognitive Control. In *Key readings in cognition. Cognitive psychology: Key readings* (pp. 205–223). Psychology Press.

Pribram, K., H. (1973). The primate frontal cortex—Executive of the brain. In K. K. H. Pribram & A. R. Luria (Eds.). In *Psychophysiology of the Frontal Lobe* (pp. 293 - 314). Academic Press.

Raffard, S., Bayard, S., Gely-Nargeot, M.-C., Capdevielle, D., Maggi, M., Barbotte, E., Morris, D., & Boulenger, J.-P. (2009). Insight and executive functioning in schizophrenia: A multidimensional approach. *Psychiatry Research, 167*(3), 239-250. https://doi.org/http://dx.doi.org/10.1016/j.psychres.2008.04.018

Raichle, M. E. (2015). The brain's default mode network. *Annu Rev Neurosci, 38*, 433-447. https://doi.org/10.1146/annurev-neuro-071013-014030

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, 98*(2), 676-682. https://doi.org/10.1073/pnas.98.2.676

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nat Rev Neurosci, 4*(3), 193-202. https://doi.org/10.1038/nrn1052

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *7*(1), 4-22. https://doi.org/10.1016/j.cogsys.2005.07.002

Reitan, R. M. (1955). The relation of the trail making test to organic brain damage. *J Consult Psychol, 19*(5), 393-394. https://doi.org/10.1037/h0044509

Ropovik, I. (2014). Do executive functions predict the ability to learn problem-solving principles? *Intelligence, 44*, 64-74. https://doi.org/10.1016/j.intell.2014.03.002

Ross, T., Calhoun, E., Cox, T., Wenner, C., Kono, W., & Pleasant, M. (2007). The reliability and validity of qualitative scores for the Controlled Oral Word Association Test. *Archives of Clinical Neuropsychology, 22*(4), 475-488. https://doi.org/10.1016/j.acn.2007.01.026

Royall, D. R., Lauterbach, E. C., Kaufer, D., Malloy, P., Coburn, K. L., & Black, K. J.. (2007). The Cognitive Correlates of Functional Status: A Review From the Committee on Research of the American Neuropsychiatric Association. The Journal of Neuropsychiatry and Clinical Neurosciences, 19(3), 249–265. https://doi.org/10.1176/jnp.2007.19.3.249

Royer, F. L., Gilmore, G. C., & Gruhn, J. J. (1984). Stimulus parameters that produce age differences in block design performance. *Journal of Clinical Psychology, 40*(6), 1474-1485. https://doi.org/10.1002/1097-4679

Rozencwajg, P., & Corroyer, D. (2001). Strategy development in a block design task. *Intelligence, 30*(1), 1-25. https://doi.org/10.1016/s0160-2896

Ruge, H., & Wolfensteller, U. (2010). Rapid Formation of Pragmatic Rule Representations in the Human Brain during Instruction-Based Learning. *Cerebral Cortex, 20*(7), 1656-1667. https://doi.org/10.1093/cercor/bhp228

Salnaitis, C. L., Baker, C. A., Holland, J., & Welsh, M. (2011). Differentiating Tower of Hanoi Performance: Interactive Effects of Psychopathic Tendencies, Impulsive Response Styles, and Modality [Article]. *Applied Neuropsychology, 18*(1), 37-46. https://doi.org/10.1080/09084282.2010.523381

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence, 39*(4), 222-232. https://doi.org/10.1016/j.intell.2011.03.001

Saracco, S. (2017). *Plato and Intellectual Development: A New Theoretical Framework Emphasising the Higher-Order Pedagogy of the Platonic Dialogues*. Springer International Publishing.

Sattler J. M. Ryan J. J. (2009). Assessment with the WAIS-IV. San Antonio, TX: Psychological Corporation.

Schmidt, J. R., & Lemercier, C. (2019). Context-specific proportion congruent effects: Compound-cue contingency learning in disguise. *Quarterly Journal of Experimental Psychology, 72*(5), 1119-1130. https://doi.org/10.1177/1747021818787155

Schneider, D. W., & Logan, G. D. (2006). Hierarchical control of cognitive processes: Switching tasks in sequences. *Journal of Experimental Psychology: General, 135*(4), 623-640. https://doi.org/10.1037/0096-3445.135.4.623

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: behavior, theory, and biological mechanisms. *Cognitive Science, 27*(3), 525-559. https://doi.org/10.1207/s15516709cog2703_8

Sebastian, A., Gerdes, B., Feige, B., Klöppel, S., Lange, T., Philipsen, A., Tebartz van Elst, L., Lieb, K., & Tüscher, O. (2012). Neural correlates of interference inhibition, action withholding and action cancelation in adult ADHD. *Psychiatry Research, 202*(2), 132-141. https://doi.org/10.1016/j.pscychresns.2012.02.010

Servant, M., Cassey, P., Woodman, G. F., & Logan, G. D. (2018). Neural bases of automaticity. *J Exp Psychol Learn Mem Cogn, 44*(3), 440-464. https://doi.org/10.1037/xlm0000454

Shallice, T. (2002). Fractionation of the supervisory system. In *Principles of frontal lobe function* (pp. 261-277). Oxford University Press.

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*. https://doi.org/10.3389/fpsyg.2014.00772

Sheu, Y.-S., & Courtney, S. M. (2016). A neural mechanism of cognitive control for resolving conflict between abstract task rules. *Cortex, 85*, 13-24. https://doi.org/10.1016/j.cortex.2016.09.018

Shiffrin, R. M., & Schneider, W. E. (1977). Controlled and automatic human information processing: Perceptual learning, automatic attending, and a general theory. *Psychol Rev, 84*(2), 128-190. https://doi.org/10.1037/0033-295X.84.1.1

Siman-Tov, T., Granot, R. Y., Shany, O., Singer, N., Hendler, T., & Gordon, C. R. (2019). Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. *Neuroscience & Biobehavioral Reviews, 105*, 262-275. https://doi.org/10.1016/j.neubiorev.2019.08.012

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*(3), 161-168. https://doi.org/10.1016/j.tins.2004.01.006

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *22*(1), 1-13. https://doi.org/10.3758/bf03202756

Snyder, H. R., Miyake, A., & Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: Bridging the gap between clinical and cognitive approaches [Review]. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00328

Sohlberg, M. M., & Mateer, C. A. (2001). Improving attention and managing attentional problems. Adapting rehabilitation techniques to adults with ADD. *Ann N Y Acad Sci, 931*, 359-375.

Spitz, H. H., Webster, N. A., & Borys, S. V. (1982). Further studies of the Tower of Hanoi problem-solving performance of retarded young adults and nonretarded children. *Developmental Psychology, 18*(6), 922-930. https://doi.org/10.1037/0012-1649.18.6.922

Spooner, D., & Pachana, N. (2006). Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology, 21*(4), 327-337. https://doi.org/10.1016/j.acn.2006.04.004

Spreen, O., Spreen, P. P. O., Strauss, E., & Strauss, P. P. E. (1998). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. Oxford University Press, USA.

Steinbeis, N., & Crone, E. A. (2016). The link between cognitive control and decision-making across child and adolescent development. *Current Opinion in Behavioral Sciences, 10*, 28-32. https://doi.org/10.1016/j.cobeha.2016.04.009

Steinborn, M. B., Langner, R., Flehmig, H. C., & Huestegge, L. (2018). Methodology of performance scoring in the d2 sustained-attention test: Cumulative-reliability functions and practical guidelines. *Psychol Assess, 30*(3), 339-357. https://doi.org/10.1037/pas0000482

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

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

Stuhr, C., Hughes, C. M. L., & Stöckel, T. (2018). Task-specific and variability-driven activation of cognitive control processes during motor performance. *Scientific Reports, 8*(1), 10811. https://doi.org/10.1038/s41598-018-29007-3

Stuss,D. T. (2011). Functions of the Frontal Lobes: Relation to Executive Functions. *Journal of the International Neuropsychological Society, 17*(05), 759-765. https://doi.org/10.1017/S1355617711000695

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychol Bull, 95*(1), 3-28. https://doi.org/10.1037/0033-2909.95.1.

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A Multidisciplinary Approach to Anterior Attentional Functionsa. *Ann N Y Acad Sci, 769*(1), 191-212. https://doi.org/10.1111/j.1749-6632.1995.tb38140.x

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *J Exp Psychol Hum Percept Perform, 13*(1), 89-103. https://doi.org/10.1037//0096-1523.13.1.89

Surrey, C., Dreisbach, G., & Fischer, R. (2017). Context-Specific Adjustment of Cognitive Control: Transfer of Adaptive Control Sets. *Quarterly Journal of Experimental Psychology, 70*(11), 2386-2401. https://doi.org/10.1080/17470218.2016.1239748

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Pearson Education.

Taylor, S. D., Bagozzi, R. P., Gaither, C. A., & Jamerson, K. A. (2006). The bases of goal setting in the self-regulation of hypertension. *J Health Psychol, 11*(1), 141-162. https://doi.org/10.1177/1359105306058869

Terry, L., & Kelley, K. (2011). Sample size planning for composite reliability coefficients: Accuracy in parameter estimation via narrow confidence intervals. *British Journal of Mathematical and Statistical Psychology, 65*, 371-401. https://doi.org/10.1111/j.2044-8317.2011.02030.x

Testa, R., Bennett, P., & Ponsford, J. (2012). Factor analysis of nineteen executive function tests in a healthy adult population. *Arch Clin Neuropsychol, 27*(2), 213-224. https://doi.org/10.1093/arclin/acr112

Tombaugh, T. (1999). Normative Data Stratified by Age and Education for Two Measures of Verbal Fluency FAS and Animal Naming. *Archives of Clinical Neuropsychology, 14*(2), 167-177. https://doi.org/10.1016/s0887-6177(97)00095-4

Tucha, L., Aschenbrenner, S., Koerts, J., & Lange, K. W. (2012). The Five-Point Test: Reliability, Validity and Normative Data for Children and Adults. *PLoS ONE, 7*(9), e46080. https://doi.org/10.1371/journal.pone.0046080

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

Tuokko, H., & Hadjistavropoulos, T. (2014). *An Assessment Guide To Geriatric Neuropsychology*. Taylor & Francis.

Uleman, J. S., & Bargh, J. A. (1989). *Unintended Thought*. Guilford Publications.

Unterrainer, J. M., Rahm, B., Kaller, C. P., Ruff, C. C., Spreer, J., Krause, B. J., Schwarzwald, R., Hautzel, H., & Halsband, U. (2004). When Planning Fails: Individual Differences and Error-related Brain Activity in Problem Solving. *Cerebral Cortex, 14*(12), 1390-1397. https://doi.org/10.1093/cercor/bhh100

Vaidya, C. J., Bunge, S. A., Dudukovic, N. M., Zalecki, C. A., Elliott, G. R., & Gabrieli, J. D. E. (2005). Altered neural substrates of cognitive control in childhood ADHD: evidence from functional magnetic resonance imaging. *Am J Psychiatry, 162*(9), 1605-1613. https://doi.org/10.1176/appi.ajp.162.9.1605

van den Heuvel, O. A., Groenewegen, H. J., Barkhof, F., Lazeron, R. H. C., van Dyck, R., & Veltman, D. J. (2003). Frontostriatal system in planning complexity: a parametric functional magnetic resonance version of tower of london task. *Neuroimage, 18*(2), 367-374. https://doi.org/10.1016/s1053-8119(02)00010-1

Vasquez, B. P., Binns, M. A., & Anderson, N. D. (2016). Staying on Task: Age-Related Changes in the Relationship Between Executive Functioning and Response Time

Consistency. *J Gerontol B Psychol Sci Soc Sci, 71*(2), 189-200. https://doi.org/10.1093/geronb/gbu140

Venegas, M. J., & Mansur, L. L. (2011). Verbal fluency: Effect of time on item generation. *Dementia & Neuropsychologia, 5*(2), 104-107. https://doi.org/10.1590/S1980-57642011DN05020008

Verbruggen, F., McLaren, I. P. L., & Chambers, C. D. (2014). Banishing the Control Homunculi in Studies of Action Control and Behavior Change. *Perspectives on Psychological Science, 9*(5), 497-524. https://doi.org/10.1177/1745691614526414

Verbruggen, F., McLaren, R., Pereg, M., & Meiran, N. (2018). Structure and Implementation of Novel Task Rules: A Cross-Sectional Developmental Study. *Psychological science, 29*(7), 1113-1125. https://doi.org/10.1177/0956797618755322

Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2008). Making choices impairs subsequent self-control: a limited-resource account of decision making, self-regulation, and active initiative. *J Pers Soc Psychol, 94*(5), 883-898. https://doi.org/10.1037/0022-3514.94.5.883

Wallesch, C. W., Curio, N., Galazky, I., Jost, S., & Synowitz, H. (2001). The neuropsychology of blunt head injury in the early postacute stage: effects of focal lesions and diffuse axonal injury. *J Neurotrauma, 18*(1), 11-20. https://doi.org/10.1089/089771501750055730

Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J., & Wagner, A. D. (2014). Frontoparietal Representations of Task Context Support the Flexible Control of Goal-Directed Cognition. *The Journal of Neuroscience, 34*(32), 10743. https://doi.org/10.1523/JNEUROSCI.5282-13.2014

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale - Fourth Edition [Measuremen Instrument]*. Pearson Clinical Assessment.

Wechsler, D. (2010). *Wechsler Abbreviated Scale of Intelligence - Second Edition [Measurement Instrument].* Pearson Clinical Assessment.

Welsh, M. C., & Huizinga, M. (2005). Tower of Hanoi disk-transfer task: Influences of strategy knowledge and learning on performance. *Learning and Individual Differences, 15*(4), 283-298. https://doi.org/http://doi.org/10.1016/j.lindif.2005.05.002

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

Wenke, D., De Houwer, J., De Winne, J., & Liefooghe, B. (2015). Learning through instructions vs. learning through practice: flanker congruency effects from instructed and applied S-R mappings. *Psychological Research, 79*(6), 899-912. https://doi.org/10.1007/s00426-014-0621-1

Westerhausen, R., Kompus, K., & Hugdahl, K. (2011). Impaired cognitive inhibition in schizophrenia: A meta-analysis of the Stroop interference effect. *Schizophrenia Research, 133*(1), 172-181. https://doi.org/https://doi.org/10.1016/j.schres.2011.08.025

Wheatley, T. P., & Wegner, D. M. (2001). Automaticity in Action. In Smelser N, J., & Baltes P, B (pp. 991-993), *International encyclopedia of the social and behavioral sciences.*

Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefcient for multidimensional measures: A pedagogical illustration. *Review of Psychology, 21*(2), 111-121.

Wiebe, S., & Karbach, J. (2018). *Executive Function: Development Across the Life Span* (1st ed.). Routledge.

Wilde, M. H., & Garvin, S. (2007). A concept analysis of self-monitoring. *Journal of Advanced Nursing, 57*(3), 339-350. https://doi.org/10.1111/j.1365-2648.2006.04089.x

Wirzberger, M., Herms, R., Esmaeili Bijarsari, S., Eibl, M., & Rey, G. D. (2018). Schema-related cognitive load influences performance, speech, and physiology in a dual-task setting: A continuous multi-measure approach. *Cognitive Research: Principles and Implications, 3*(1). https://doi.org/10.1186/s41235-018-0138-z

Yoshida, W., Funakoshi, W., & Ishii, S. (2010). Hierarchical rule switching in prefrontal cortex. *Neuroimage, 50*(1), 314-322. https://doi.org/10.1016/j.neuroimage.2009.12.017

Yoshida, W., & Ishii, S. (2006). Resolution of Uncertainty in Prefrontal Cortex. *Neuron, 50*(5), 781-789. https://doi.org/10.1016/j.neuron.2006.05.006

Yuan, K.-H., Bentler, P. M., & Zhang, W. (2005). *The Effect of Skewness and Kurtosis on Mean and Covariance Structure Analysis: The Univariate Case and Its Multivariate Implication*. Sage Publications.

Zalonis, I., Christidi, F., Bonakis, A., Kararizou, E., Triantafyllou, N. I., Paraskevas, G., Kapaki, E., & Vasilopoulos, D. (2009). The Stroop Effect in Greek Healthy Population: Normative Data for the Stroop Neuropsychological Screening Test. *Archives of Clinical Neuropsychology, 24*(1), 81-88. https://doi.org/10.1093/arclin/acp011

Zink, N., Stock, A.-K., Vahid, A., & Beste, C. (2018). On the Neurophysiological Mechanisms Underlying the Adaptability to Varying Cognitive Control Demands. *Frontiers in Human Neuroscience, 12*. https://doi.org/10.3389/fnhum.2018.00411

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition, 56*(3), 286-292. https://doi.org/http://doi.org/10.1016/j.bandc.2004.07.003

Appendix A Promotional Flyer Distributed To Engage Study Participation

VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

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 18-50 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:

Adam Bromage on Adam.Bromage@live.vu.edu.au

Ph: 0414 229 314

This study is conducted under the supervision of Dr Michelle Ball and Dr Emra Suleyman.

Appendix B Information to Participants Form

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 student Anique Muttiah.

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 EFs, 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 on 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) or 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
- Anique Muttiah – Honours investigating whether EF tests can be used to predict procrastination and perfectionist tendencies.

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 student Anique Muttiah. 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
michelle.ball@vu.edu.au
9919 2536

Emra Suleyman
emra.suleyman@vu.edu.au
9919 2397

Jessica Burlak
Jessica.burlak@vu.edu.au
0411 575 176

Adam Bromage
adam.bromage@vu.edu.au
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 C Consent Form for Participants Involved in Research

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 D Alternative Test Administration Orders

Assessment Version			
Version A	Version B	Version C	Version D
Block Design (WASI)	**Block Design (WASI)**	**Block Design (WASI)**	**Block Design (WASI)**
Vocabulary (WASI)	**Vocabulary (WASI)**	**Vocabulary (WASI)**	**Vocabulary (WASI)**
TMT – A	**TOH**	**5- Point Test**	Digit Span FWD
TMT – B	**TMT – A**	**Stroop Test**	**Digit Span BWD**
5-Point Test	**TMT – B**	**TMT – A**	Visual Span FWD
TOH	Picture Arrangement (WAIS III)	**TMT – B**	**Visual Span BWD**
Test of D2	**5-Point Test**	ROCF Copy Trial	**Map Search (TEA)**
Digit Span FWD	WCST	**FAS Test**	Elevator Counting (TEA)
Digit Span BWD	Digit Span FWD	ROCF Delay Trial	**Visual Elevator (TEA)**
Visual Span FWD	**Digit Span BWD**	**Test of D2**	**Elevator Counting Reversal (TEA)**
Visual Span BWD	Visual Span FWD	Fluency (Animals)	Telephone Search (TEA)
Stroop Test	**Visual Span BWD**	Rule Shift (BADS)	Telephone Search While Counting (TEA)
ROCF Copy Trial	Rule Shift (BADS)	Key Search (BADS)	Lottery (TEA)
FAS Test	Key Search (BADS)	Zoo Map 1 (BADS)	**Austin Maze**
ROCF Delay Trial	Zoo Map 1 (BADS)	Zoo Map 2 (BADS)	Picture Arrangement (WAIS III)
Fluency (Animals)	Zoo Map 2 (BADS)	*BREAK*	**Five-Point Test**
Austin Maze	*BREAK*	Digit Span FWD	*BREAK*
BREAK	**Map Search (TEA)**	**Digit Span BWD**	**TOH**
Picture Arrangement (WAIS III)	Elevator Counting (TEA)	Visual Span FWD	Fluency (Animals)
WCST	**Visual Elevator (TEA)**	**Visual Span BWD**	WCST
Rule Shift (BADS)	**Elevator Counting Reversal (TEA)**	**Map Search (TEA)**	**TMT – A**
Key Search (BADS)	Telephone Search (TEA)	Elevator Counting (TEA)	**TMT – B**
Zoo Map 1 (BADS)	Telephone Search While Counting (TEA)	**Visual Elevator (TEA)**	**Stroop Test**
Zoo Map 2 (BADS)	Lottery (TEA)	**Elevator Counting Reversal (TEA)**	ROCF Copy Trial
Map Search (TEA)	**Stroop Test**	Telephone Search (TEA)	**FAS Test**
Elevator Counting (TEA)	**Austin Maze**	Telephone Search While Counting (TEA)	ROCF Delay Trial
Visual Elevator (TEA)	ROCF Copy Trial	Lottery (TEA)	**Test of D2**
Elevator Counting Reversal (TEA)	**FAS Test**	**Austin Maze**	Rule Shift (BADS)
Telephone Search (TEA)	ROCF Delay Trial	Picture Arrangement (WAIS III)	Key Search (BADS)
Telephone Search While Counting (TEA)	**Test of D2**	**TOH**	Zoo Map 1 (BADS)
Lottery (TEA)	Fluency (Animals)	WCST	Zoo Map 2 (BADS)

Note. Tests in bold type face were selected for inclusion within the current study. FWD = Forwards; BWD = Backwards; ROCF = Rey-Osterrieth Complex Figure Test; WCST = Wisconsin Card Sorting Test; BADS = Behavioural Assessment of the Dysexecutive Syndrome.

Appendix E Supporting Matrices for Exploratory Factor Analysis of Tower of Hanoi

Table E1

Factor Correlation Matrix with Direct Oblimin Rotation for Tower of Hanoi.

Factor	1	2	3	4
1	1.000	-.073	.026	.074
2	-.073	1.000	-.080	-.195
3	.026	-.080	1.000	.165
4	.074	-.195	.165	1.000

Table E2

Standardised Residual Covariances for Tower of Hanoi

	Trial 11	Trial 8	Trial 7	Trial 5	Trial 3
Trial 11	.000				
Trial 8	-.209	.000			
Trial 7	.174	-.141	.000		
Trial 5	.298	.363	-.722	.000	
Trial 3	-.157	-.278	1.229	-.295	.000

Table E3

Factor Score Weights for TOH

	Trial 11	Trial 8	Trial 7	Trial 5	Trial 3
TOH	.009	.125	.046	.070	.047

Appendix F Demand Classification System Criteria and Scoring Sheets for all Tests.

Table F1

DCS Criteria for Trials 5 – 9 of the Block Design Test

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
(1)	Abstraction					
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established via the provision of a design. The image provides a visual configuration of the design that requires minimal formulation of sub-goals.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	The evaluation of the physical state of the colour blocks is required to enable accurate configuration and placement in accordance to the prescribed design.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The physical dimensions of two different solid colour sides and one split colour side of each block must be considered and relationships formed to establish the correct combination of colour sides to achieve the prescribed design.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The operational parameters of the trial trials are stable comprising of four blocks requiring manipulation of block into a square design configuration. Configurations change between each trial is provided by the design book.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Task requires selection of four correct block colours (e.g. solid colour side, vs. split red & white sides) amongst four blocks containing 16 possible sides.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are to establish the response and speed requirements for each trial. Rules are minimal and not uniquely exclusive to this task environment alone (e.g. do not rotate the stimulus book at any time).
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly, and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	**6.F**	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	The task requires manipulation of four blocks by rotating and positioning them to match the prescribed design.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	The task requires the fundamental knowledge that separate shapes when placed together can collectively establish a global design/pattern.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	The task does not require the retention of pattern or configuration design for successful completion of each trial.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 5 – 9 within the Block Design Test	Raw performance scores

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Criteria: 3

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 3 Maximum = 5	**Total score for (C) =** 4 Maximum = 10

Total Score for Complexity (S) + (C) 7

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 3 Maximum = 3	**Total score for (N) =** 0 Maximum = 6

Total Score for Novelty (F) + (N) 3

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F2

DCS criteria for Trials 10 & of the Block Design Test

ID	Demand Factors		Demand Criteria			Task Specific Information & Requirements
			(S) Simple		(C) Complex	
(1)	Abstraction	1.S		1.C		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	1T.C	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established via the provision of a design. The image provides a visual configuration of the design but requires the formulation of sub-goals to achieve various smaller shape designs that comprise the superordinate design.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	1P.C	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	The evaluation of the physical state of the colour blocks is required to enable accurate configuration and placement in accordance to the prescribed design.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	1R.C	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The physical dimensions of two different solid colour sides and one split colour side of each block must be considered and relationships formed to establish the correct combination of colour sides to achieve the prescribed design.
(2)	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue	2.C	**The operational parameters of the task change over the course of administration. Changes may not be alerted via an exogenous cue**	The operational parameters of the trials change requiring use of an additional 5 blocks and the exclusive use of split-colour blocks to establish the prescribed design during Trial 11.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	3.C	**Many alternative response options are available –**	Task requires selection of 9 correct block colours (e.g. solid colour side, vs. split red & white sides) amongst four blocks containing 36 sides.
(4)	Instructions and rules	4.S	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are to establish the response and speed requirements for each trial. Rules are minimal and not uniquely exclusive to this task environment alone (e.g. do not rotate the stimulus book at any time).
(5)	Dual Nature	5.S	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	The task requires manipulation of nine blocks by rotating and positioning them to match the prescribed design.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	7.N	**Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas**	The task requires the adaptation of fundamental knowledge on shapes to establish a complex design that includes various alternative and novelty shapes .
(8)	Episodic Demands	8.F	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	The task does not require the retention of pattern or configuration for successful completion of each trial.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 10 & 11within the Block Design Test	Raw performance scores

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Criteria: 3

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
	Total Score for each Demand Criteria	**Total score for (S) =** 2 Maximum = 5	**Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
	Total Score for each Demand Criteria	**Total score for (F) =** 2 Maximum = 3	**Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Familiar

Table F3

DCS criteria for Trials 12 & 13 of the Block Design Test

ID	Demand Criteria		Demand Classifications: (S) Simple		Demand Classifications: (C) Complex	Task Specific Information & Requirements
(1)	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established during Trial 12 but not Trial 11, with the removal of the border guide. The image provides a visual configuration of the design and but requires the formulation of sub-goals to achieve various smaller shape designs that comprise the superordinate design.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	The evaluation of the physical state of the colour blocks is required to enable accurate configuration and placement in accordance to the prescribed design.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The physical dimensions of two different solid colour sides and one split colour side of each block must be considered and relationships formed to establish the correct combination of colour sides to achieve the prescribed design.
(2)	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue	**2.C**	**The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue**	The operational parameters of the trials change with the removal of the border guide from the design, and complete use of split-sided blocks to establish the prescribed design during Trial 1 12 &13.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Task requires selection of 9 correct block colours (e.g. solid colour side, vs. split red & white sides) amongst four blocks containing 36 sides.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are to establish the response and speed requirements for each trial. Rules are minimal and not uniquely exclusive to this task environment alone (e.g. do not rotate the stimulus book at any time).
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	The task requires the rotating and positioning of each block to match an ambiguous prescribed design.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas**	The task requires the adaptation of fundamental knowledge on shapes to establish a complex design that includes various alternative and novelty shapes.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	The task requires experience with Trial 12 to be carried to Trial 13. Successful transfer of this S→R from Trial 12 due to diamond configuration will improve application of the diamond configuration to the novel ambiguity presented in Trial 13.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 12 & 13 within the Block Design Test	Raw performance scores

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Criteria: 3

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 2 Maximum = 5 — **Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 0 Maximum = 3 — **Total score for (N) =** 6 Maximum = 6

Total Score for Novelty (F) + (N) 6

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Novel

Table F4
DCS Criteria for Trials 3, 4, & 5 of the Austin Maze

ID	Demand Factors		Demand Criteria: (S) Simple		(C) Complex	Task Specific Information & Requirements
(1)	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established via the provision of start and end tiles. Sub-goals are required to be established to navigate each direction in the path sequence of maze in conjunction with feedback from the apparatus.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	Task requires the evaluation of the physical position of each tile within the maze to enable an appropriate response.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The properties of the maze are concrete with minimal relationships between dimensions of the maze stimuli requiring consideration to response.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The maze does not change in configuration.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Correct tiles are hidden amongst numerous visually identical tiles.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Multiple rules are enforced that are exclusive to the task (e.g. do not travel backwards, in a diagonal direction, or skip a tile).
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Each trial requires the repetition or correction of action via feedback from previous S→R experience to be integrated into each new trial.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas**	The task requires the acquisition of task specific schemas to learn the hidden path via S→R.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	The task requires the transfer of experience with each previous trial to successfully until the hidden path is learned.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 3, 4, & 5 within the Austin Maze	Number of correctly identified tiles

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Criteria: 2

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 2 Maximum = 5	**Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 0 Maximum = 3	**Total score for (N) =** 6 Maximum = 6

Total Score for Novelty (F) + (N) 6

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Novel

Table F5

DCS Criteria for Trials 7, 8 & 9 of the Austin Maze

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Information & Requirements
(1)	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established via the provision of start and end tiles. Sub-goals are required to be established to navigate each direction in the path sequence of maze in conjunction with feedback from the apparatus.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	Task requires the evaluation of the physical position of each tile within the maze to enable an appropriate response.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The properties of the maze are concrete with minimal relationships between dimensions of the maze stimuli requiring consideration to response.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The maze does not change in configuration.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Correct tiles are displayed amongst numerous visually identical tiles.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Multiple rules are enforced that are exclusive to the task environment (e.g. do not travel backwards, in a diagonal direction, or skip a tile).
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Each trial requires the repetition or correction of action via feedback from previous S→R experience to be integrated into each new trial.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	The task required the direct application of implicit task specific schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	The task required the application of a well-practiced path.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 7, 8 & 9 within the Austin Maze	Number of correctly identified tiles

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 2

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 2 Maximum = 5 — **Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 2 Maximum = 3 — **Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION — **Global Complexity Demand** + **Global Novelty Demand** = Complex & Familiar

Table F6

DCS Criteria for 0-15-seconds of the FAS Test

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Information & Requirements
(1)	Abstraction	**1.S**		2.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established via the exogenous provision of each letter.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Task provides an exogenous guide in the form of a starting letter for each class of words, simple rules that links S→R (e.g. new words, none repeated).
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The properties of the task are concrete with minimal relationships between letters or responses requiring consideration in order to respond.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The change to each letter is altered by an exogenous cue.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	All words within the dictionary under each letter that follow the task rules are possible available responses.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Multiple rules that govern the use of words that are proper names or have a suffix. The application of this rule involved basic understanding of language that is not unique to this task.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	**6.F**	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	Task requires application of vocabulary and semantics that form the basis of everyday language accessed via semantic memory.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires that application of everyday language schemas for a prescribed letter and general language rules.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	The task requires the retention of previously used words to avoid repetition of words.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
0-15-seconds of the FAS Test	Total number of correct words

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 4 Maximum = 5 — **Total score for (C) =** 2 Maximum = 10

Total Score for Complexity (S) + (C) 6

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 2 Maximum = 3 — **Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F7

DCS Criteria for 16- 60 seconds of the FAS Test

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Information & Requirements
(1)	Abstraction	**1.S**		1.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established via the exogenous provision of each letter
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Task provides an exogenous guide in the form of a starting letter for each class of words, simple rules that links S→R (e.g. new words, none repeated).
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The properties of the task are concrete with minimal relationships between letters or responses requiring consideration in order to respond.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The change to each letter is altered by an exogenous cue.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	All words that adhere to the instructions and rules of the task that fall under each letter within the dictionary are possible response options available.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Multiple rules that govern the use of words that are proper names or have a suffix. This rule is not uniquely available to this task as forms part of everyday language.
(5)	Dual Nature	5.S	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	6.N	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires vocabulary and semantics that form the basis of everyday language. As immediate recall of words from semantic memory are exhausted, explicit recall and adaptation of new words is required to complete the task.
(7)	Schematic Demands	7.F	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires that application of everyday language schemas for a prescribed letter and general language rules.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	8.N	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	The task requires the retention of previously used words to avoid repletion or words.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
16-60-seconds of the FAS Test	Total number of correct words

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 4 Maximum = 5 — **Total score for (C) =** 2 Maximum = 10

Total Score for Complexity (S) + (C) 6

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 1 Maximum = 3 — **Total score for (N) =** 4 Maximum = 6

Total Score for Novelty (F) + (N) 5

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION — **Global Complexity Demand** + **Global Novelty Demand** = Simple & Novel

Table F8
DCS Criteria for Trials 3,5,7,8 and 11 of the Tower of Hanoi

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
1.	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established during instructional period. A series of sub goal are required to formed to establish the correct sequence of movements in order to complete the task with minimal movements.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	Higher-order abstract policies are required to evaluate the prescribed goal state and the sequence of movements required from the start state.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	Relationships are required to be considered between the size of the discs, required movements, and the goal configuration to the tower.
2.	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue	**2.C**	**The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue**	The parameters of the task change with the addition of a fourth disc.
3.	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Many alternative movement combinations are available.
4.	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Multiple instructions and rules are given that govern the execution of disc movement that are unique to the task environment.
5.	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	The task only requires completion of a single tower trial at one time.
			(F) Familiar		(N) Novel	
6.	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires the application of physical movements and visual monitoring to a series of explicit task specific movements.
7.	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires the repetition of fundamental schemas of disc size and physical movements. The schemas are implicit to the TOH across all trials and unchanging.
8.	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Task does not require the transfer of configurations to other trials.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Tower of Hanoi Trials 3, 5 , 7, 8, 11	Residual moves score

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 3

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 1 Maximum = 5 — **Total score for (C) =** 8 Maximum = 10

Total Score for Complexity (S) + (C) 9

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 2 Maximum = 3 — **Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Familiar

Table F9

DCS Criteria for Stroop Test – Word Trial

ID	Demand Factors		Demand Criteria (S) Simple		Demand Criteria (C) Complex	Task Specific Details
(1)	Abstraction					
	(T) Temporal	1T.S	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established via the provision of all words to be read aloud.
	(P) Policy	1P.S	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Task provides an exogenous stimulus card of words.
	(R) Relational	1R.S	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response.	The properties of the task are concrete with minimal with no relationships between the words needing to be considered.
(2)	Contextual Stability	2.S	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue.	The task is stable with no changes to the stimulus card during the trial.
(3)	Action Rules	3.S	**A minimal subset of possible response options are available**	3.C	Many alternative response options are available	All words are prescribed and made available.
(4)	Instructions and rules	4.S	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are singular that require the prescribed words to be read aloud.
(5)	Dual Nature	5.S	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Only one word is required to be read at one time.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	Task requires reading ability that forms the basis of everyday language.
(7)	Schematic Demands	7.F	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires the application of fundamental knowledge of written language.
(8)	Episodic Demands	8.F	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	No retention of words is required during the trial.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Stroop Test – Word Trial	Total time

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 5 Maximum = 5 — **Total score for (C) =** 0 Maximum = 10

Total Score for Complexity (S) + (C) 5

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 3 Maximum = 3 — **Total score for (N) =** 0 Maximum = 6

Total Score for Novelty (F) + (N) 3

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION — **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F10

DCS Criteria for Stroop Test: Colour-Word Trial

ID	Demand Criteria		Demand Classifications (S) Simple		(C) Complex	Task Specific Information & Requirements
(1)	Abstraction	**1.S**		1.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are established via the provision of all items to be read aloud.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	A higher order policy is established that requires evaluation between the written words and the colour that they printed in order to select an appropriate response.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	Relationships between the two interdependent dimensions (written words and colour ink) must be considered in order to select correct. response
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the stimulus card during the trial.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	All stimuli are prescribed and made available, however competing response alternatives are present between either the naming of the ink colour, or the name that of written word.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Application of instruction required to be applied to form a rule unique task environment due to incongruently displayed colour words and colour they are printed in.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Only one word is required to be read at one time.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires automatic reading abilities to be to be inhibited in place of colour naming that is specific to the incongruency of the stimuli.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires use of explicit of task S→R to create new task specific schemas**	Task requires the application of fundamental knowledge of written language and colour naming to be utilised in a new task specific environment that requires dominancy over the colour naming schemas over reading schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	No retention of words is required during the trial.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Stroop Test: Colour-Word Trial	Total time

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 2

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 2 Maximum = 5 — **Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 1 Maximum = 3 — **Total score for (N) =** 4 Maximum = 6

Total Score for Novelty (F) + (N) 5

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Novel

Table F11

DCS Criteria for TMT-A

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
(1)	Abstraction	**1.S**		1.C		
(T)	Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established via the exogenous provision of all numbers with the numerical sequence being cued by the availability of the next number.
(P)	Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Task provides an exogenous stimulus card of all numbers with no higher order policy determining response.
(R)	Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The properties of the task are concrete with only relationships between the numerical order of the numbers needing to be considered.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the stimulus card during the trial.
(3)	Action Rules	**3.S**	**A minimal subset of possible response options are available**	3.C	Many alternative response options are available	A small set of numbers is prescribed and made available.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are singular with the pen to remain on the page and the numbers to be connected in numerical order.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Only one sequence of numbers is required to be connected at one time.
			(F) Familiar		(N) Novel	
(6)	Automaticity	**6.F**	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	Task requires the knowledge of drawing a continuous line, reading, and counting.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires use of explicit of task S→R to create new task specific schemas	Task requires the understanding and application of numerical order and search schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	No retention of numbers or tracing is required during the trial.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
TMT-A	Total time

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
	Total Score for each Demand Criteria	**Total score for (S) =** 5 Maximum = 5	**Total score for (C) =** 0 Maximum = 10

Total Score for Complexity (S) + (C) 5

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
	Total Score for each Demand Criteria	**Total score for (F) =** 3 Maximum = 3	**Total score for (N) =** 0 Maximum = 6

Total Score for Novelty (F) + (N) 3

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F12

DCS Criteria for TMT-B

ID	Demand Criteria		(S) Simple		(C) Complex	Task Specific Details
(1)	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established via the exogenous provision of all numbers and letters Sub-goals are required to be formulated that integrate numbers and letters in their correct sequence in order to achieve the superordinate goal.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Task provides an exogenous stimulus card of numbers and letters.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The properties of the task are concrete but relationships between the numerical order and alphabetical order need to be considered in order to select a correct response.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the stimulus card during the trial.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	A larger subset of numbers and letters is made available.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are singular with the pen to remain on the page and the numbers and letters to be connected in alternating between alphabetical and numerical order.
(5)	Dual Nature	5.S	Task requires completion of a singular task	**5.C**	**Task encompasses a dual or multi-task requirement**	A task of sequencing between alphabetical and numerical order is conducted.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires the use of drawing a continuous line, reading, and counting abilities to be integrated into a new a sequence of number-letter-number-letter responses.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires use of explicit of task S→R to create new task specific schemas	Task requires the understanding and application of numerical and alphabetical orders, and line tracing for letter-number sequencing.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Retention of place in the number sequence whilst dealing with the letter sequence and vice versa.is required.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
TMT-B	Total Time

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of **Complex** (T) (P) and (R) Classifications: 2

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 2 Maximum = 5	**Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 2 Maximum = 3	**Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Familiar

Table F13
DCS Criteria for 0-60seconds of the 5-point Test

ID	Demand Factors		Demand Criteria (S) Simple		Demand Criteria (C) Complex	Task Specific Details
1	Abstraction					
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established during the instructional period. Each figure may constitute a sub-goal, however each sub-goal exists independently and does not contribute towards a global figure.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	The task provides guides for the figure to be completed within which provide a link between the figure and the response available.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	Relationships between the stimulus sheet and the intended figures is required to be considered in order to create and produce an accurate response.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the stimulus card during the trial.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	A large subset of possible figures are available and must require evaluation before responding.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Instructions are given with multiple rules that govern the creation of figures that are exclusive to the task environment.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement.	Each figure is completed individually.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires fundamental knowledge of shapes to be applied to establish series of task specific configurations.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires the direct application of implicit schemas for shapes and design schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Previously drawn figures are within view and within close visual range of the participant reducing the need to retain and transfer previous representation to memory.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
5-Point Test 0-60 seconds	Total number of correct figures

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of **Complex** (T) (P) and (R) Classifications: 1

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 3 Maximum = 5	**Total score for (C) =** 4 Maximum = 10

Total Score for Complexity (S) + (C) 7

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 2 Maximum = 3	**Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F14

DCS Criteria for 61-120seconds trial of the 5-point Test

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
(1)	Abstraction	**1.S**		1.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment.	Superordinate goals are explicitly established during instructional period. Each figure may constitute a sub-goal, however each sub-goal exists independently and does not contribute towards a global figure.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	The task provides guides for the figure to be completed within which provide a link between the figure and the response available.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	Relationships between the stimulus sheet and the intended figures is required to be considered in order to create and produce an accurate response.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the stimulus card during the trial.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	A large subset of possible figures are available to be created.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Instructions are given with multiple rules that govern the creation of figures that are exclusive to the task environment.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each figure is completed individually.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires fundamental knowledge of shapes to be applied to establish series of task specific configurations.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires the implicit schemas for shapes and design schemas.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	Previously drawn figures are within view but are out of immediate visual range increasing the need to retain and transfer previous responses from memory.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
5-point Test 61-120 seconds trial	Total number of correct figures

Step 1	**Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 1

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2	**Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 3 Maximum = 5	**Total score for (C) =** 4 Maximum = 10

Total Score for Complexity (S) + (C) 7

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3	**Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 1 Maximum = 3	**Total score for (N) =** 4 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Novel

Table F15

DCS criteria for 0-60seconds of Map Search

ID	Demand Factors		Demands Criteria: (S) Simple		Demands Criteria: (C) Complex	Task Specific Details
(1)	Abstraction	**1.S**		1.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established during instructional period. Required actions are made explicit and prompted with the provision of cued symbols.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	A lower-order abstract policy exists where 80 exogenously provided target symbols must be found within a map.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The map does not require any relationships between the map and the prescribed symbol to be evaluated for completion.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the map or the symbol.
(3)	Action Rules	**3.S**	**A minimal subset of possible response options are available**	3.C	Many alternative response options are available	A correct target symbol is explicitly provided amongst distractor symbols.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions that set the purpose of the task are followed by an instructional-rule that governs which symbol must be selected.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	The identification of the symbol is the only task required for completion.
			(F) Familiar		(N) Novel	
(6)	Automaticity	**6.F**	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	Task requires that visual scanning and identification of a familiar symbol amongst a generic street map environment.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires only the direct application of search and respond schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Symbols are not required to be retained to memory, and the target symbol is explicitly present throughout completion.

Note: Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Map Search 0-60seconds	Total number correctly identified symbols

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 5 Maximum = 5 — **Total score for (C) =** 0 Maximum = 10

Total Score for Complexity (S) + (C) 5

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 3 Maximum = 3 — **Total score for (N) =** 0 Maximum = 6

Total Score for Novelty (F) + (N) 3

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F16

DCS Criteria for 61-120seconds of Map Search

ID	Demand Factors		Demand Criteria: (S) Simple		(C) Complex	Task Specific Details
(1)	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established during instructional period. Required actions are made explicit with the provision of a cued target symbol. Limited availability of symbols exists, requiring the additional need to adopt a search strategy.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	A higher order abstract policy must be formed to enact a response that allows for selection of amongst the few correct symbols amongst distractors.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The map does not require any relationships to be evaluated for completion.
(2)	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue	**2.C**	**The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue**	A visual change to the map occurs via the cancelling out of symbols over the course of administration, in addition to the changing of the colour of the pen at the start of the trial.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Many incorrect alternative response options are now available with correct responses becoming difficult to isolate.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions that set the purpose of the task are followed by an instructional rule that governs which symbol must be selected.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	The identification of the symbol is the only task required for completion.
			(F) Familiar		(N) Novel	
(6)	Automaticity	**6.F**	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	Task requires that visual scanning and identification of a familiar symbol amongst a generic street map environment.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires only the direct application of search and respond schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Symbols are not required to be retained to memory, and the target symbol is explicitly present throughout completion.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Map Search 61-120seconds	Total number correctly identified symbols

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 2

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2: **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria — **Total score for (S) =** 2 Maximum = 5 — **Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria — **Total score for (F) =** 3 Maximum = 3 — **Total score for (N) =** 3 Maximum = 6

Total Score for Novelty (F) + (N) 3

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Familiar

Table F17

DCS Criteria for Digit Span Backwards

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
(1)	Abstraction	1.S		1.C		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Task requires sub-goals for each digit (or a chunking of numbers) to enable the digit to be recalled in reverse to the prescribed order.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	The policy between the digits prescribed and the requirements to read them in reverse is provided exogenously.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	No independent relationships between the digits are required to be evaluated.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimal changes Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The operation parameters of the task remain stable within minimal changes via the additional of one extra digit for every two correct responses.
(3)	Action Rules	**3.S**	**A minimal subset of possible response options are available**	3.C	Many alternative response options are available	All digits required that are to be recalled are prescribed and made available.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are given that stipulate a rule that all digits must be recalled in reverse order.
(5)	Dual Nature	5.S	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Only one string of digits is required to be recalled at one time.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires ability to recall spoken words in the form of numbers to be manipulated into a reverse order.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires use of explicit of task S->R to create new task specific schemas**	Task requires the application of fundamental schemas of numbers and ordering to be applied to a reverse-order specific schema of the task.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	Retention of a prescribed string of digits is required to be transferred to the reverse rule requirement.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Digit Span – Backwards	Total number of correct trials

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of **Complex** (T) (P) and (R) Classifications: 1

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 5 Maximum = 5	**Total score for (C) =** 0 Maximum = 10

Total Score for Complexity (S) + (C) 5

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 0 Maximum = 3	**Total score for (N) =** 6 Maximum = 6

Total Score for Novelty (F) + (N) 6

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Novel

Table F18

DCS Criteria for Visual Span Backwards

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
(1)	Abstraction	1.S		1.C		
(T)	Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Task requires that sub-goals to be established for each block so that the sequence of blocks can be repeated in a reverse order.
(P)	Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	The policy between the block order prescribed and the requirements to repeat them in reverse is provided exogenously.
(R)	Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	No independent relationships between the blocks are required to be evaluated.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimal changes Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The operation parameters of the task remain stable within minimal changes via the additional of one extra block for every two correct responses.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	The correct sequence of blocks must be selected amongst visually identical block stimuli.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are given that stipulate the rule that all block sequences must be recalled in reverse.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Only one block sequence is required to be recalled at one time.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires ability to recall visual sequences and patterns to be manipulated into a reverse order.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires use of explicit of task S→R to create new task specific schemas**	Task requires the application of fundamental knowledge of sequences and ordering to be applied to a reverse-order specific schema of the task.
(8)	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	Retention of a prescribed sequence is required and transferred to the reverse rule requirement.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Visual Span – Backwards	Total number of correct trials

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 1

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Criteria	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2

Total Score for each Demand Criteria **Total score for (S) =** 4 Maximum = 5 **Total score for (C) =** 2 Maximum = 10

Total Score for Complexity (S) + (C) 6

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Criteria	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2

Total Score for each Demand Criteria **Total score for (F) =** 0 Maximum = 3 **Total score for (N) =** 6 Maximum = 6

Total Score for Novelty (F) + (N) 6

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Novel

Table F19
DCS Criteria for Test of d2

ID	Demand Factors		Demand Criteria: (S) Simple		(C) Complex	Task Specific Details
(1)	Abstraction	**1.S**		1.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established during instructional period. Required actions are made explicit and prompted with the provision of cued symbols.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	A lower-order abstract policy exists where exogenously provided target symbols must be found within the response sheet.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	Only dimensions between symbol configurations are required to be evaluated.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The task is stable with no changes to the response sheet during the trial.
(3)	Action Rules	**3.S**	**A minimal subset of possible response options are available**	3.C	Many alternative response options are available	Correct target symbols are explicitly provided amongst distractor symbols.
(4)	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions that set the purpose of the task are followed by an instructional-rule that governs which symbols must be selected.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each symbol and line of symbols is completed in a singular fashion.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires that visual scanning and identification of letters to be applied to a unique symbol not typically found.
(7)	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	Task requires only the direct application of search and respond schemas.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Symbols are not required to be retained in memory as target symbols are is provided for reference throughout completion.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Test of d2	Total number of correctly identified configurations.

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
	Total Score for each Demand Criteria	**Total score for (S) =** 5 Maximum = 5	**Total score for (C) =** 0 Maximum = 10

Total Score for Complexity (S) + (C) 5

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
	Total Score for each Demand Criteria	**Total score for (F) =** 2 Maximum = 3	**Total score for (N) =** 2 Maximum = 6

Total Score for Novelty (F) + (N) 4

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Familiar

Table F20

DCS Criteria for Visual Elevator

ID	Demand Factors		Demand Criteria: (S) Simple		(C) Complex	Task Specific Details
(1)	Abstraction	**1.S**		1.C		
	(T) Temporal	**1T.S**	**Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues**	1T.C	Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment	Superordinate goals are explicitly established during instructional period. Required actions are made explicit with the provision of a cued target symbols in the form of elevators and directional arrows.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Lower order policies are established that stipulate counting direction when exogenous arrows are present.
	(R) Relational	**1R.S**	**Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response**	1R.C	Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response	The cue for response that the elevator symbols and arrows provide remains stable, with a minimal need to attribute directionality of counting to the arrows presented.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The parameters of the task are stable with the elevator scene remaining in place, with minimal changes to the quantity of elevator images and arrows between tasks.
(3)	Action Rules	**3.S**	**A minimal subset of possible response options are available**	3.C	Many alternative response options are available	The test prescribes the available options that guide response.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Instructions and rules are introduced that govern the sequence of responses required when presented with an elevator, up, or down arrow stimuli.
(5)	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial is completed individually and singularly.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires the integration of visual scanning and counting to be applied to the task specific elevator scenario.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas**	Task requires the adaptation of counting schemas towards the formation of task specific schemas. E.g. 'if an up arrow is seen, begin to count forwards after, if a down arrow is seen begin to count backwards after'.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Reponses are not required to be retained to memory to learn new sequences.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Visual Elevator	Total timing score for correct trials

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 0

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 1

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 4 Maximum = 5	**Total score for (C) =** 2 Maximum = 10

Total Score for Complexity (S) + (C) 6

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Simple

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 1 Maximum = 3	**Total score for (N) =** 4 Maximum = 6

Total Score for Novelty (F) + (N) 5

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Simple & Novel

Table F21

DCS Criteria for ECR

ID	Demand Factors		Demand Criteria (S) Simple		Demand Criteria (C) Complex	Task Specific Details
(1)	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established during instructional period. Task requires the series of sub-goals in the form counting of forwards and backwards is response to audio cues of different frequency. These cues inform the direction of counting.
	(P) Policy	**1P.S**	**Lower-order policies contain simple imposed rules that link S→R**	1P.C	There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response	Lower-order policies are established that stipulate counting direction in response to the exogenous tones present.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The task required the continuous implicit evaluation of the correct counting sequence to the correct audio cue.
(2)	Contextual Stability	**2.S**	**The operational parameters of the task(s) condition(s) are stable with minimum changes. Any changes between trials are alerted by an exogenous cue**	2.C	The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue	The parameters of the task are stable with the elevator scene remaining in place, with minimal changes to the quantity of audio tones and arrows between tasks.
(3)	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	The test requires a different response for three different Auditory cues. A response must be chosen accurately for every cue. Alternative response may be given in error if the cue is not attributed correctly.
(4)	Instructions and rules	4.S	Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment	**4.C**	**Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment**	Instructions and rules are introduced that govern the task response required when a specific tone is heard. These instructions and rules are unique to the task environment.
(5)	Dual Nature	5.S	Task requires completion of a singular task	**5.C**	**Task encompasses a dual or multi-task requirement**	The task requires counting forwards and backwards sub-vocally in addition to the intake and updates of information via audio cues.
			(F) Familiar		(N) Novel	
(6)	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	Task requires the integration of counting sequences to be applied intermittently to task specific requirements whilst continuously receiving new information that governs the sequence.
(7)	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas**	Task requires the adaptation of counting schemas towards the formation of task specific schemas E.g. 'if hearing a high pitch tone, begin to count forwards after, if a low pitch tone is heard, begin to count backwards after, but do not count either as part of the sequence'.
(8)	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	Strings of counting are required to be held whilst direction tones intermit the sequence, however this does not need to be transferred to a new sequence trials.

Note. Text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Elevator Counting with Reversal (ECR)	Total number of correct trials

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 2

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 1 Maximum = 5	**Total score for (C) =** 8 Maximum = 10

Total Score for Complexity (S) + (C) 9

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 1 Maximum = 3	**Total score for (N) =** 4 Maximum = 6

Total Score for Novelty (F) + (N) 5

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Novel

Appendix G Matrices from Block Design Analysis

Table G1

Covariance Matrix of all Block Design Trials Analysed

Trial	BD 13	B 12	BD 11	BD 10	BD 9	BD 8	BD 7	BD 6	BD 5
BD 13	1.398								
BD 12	0.816	1.549							
BD 11	0.774	0.918	1.559						
BD 10	0.647	0.887	0.823	1.51					
BD 9	0.591	0.792	0.806	0.854	1.544				
BD 9	0.538	0.791	0.686	0.604	0.683	1.3			
BD 7	0.42	0.581	0.532	0.551	0.726	0.589	1.115		
BD 6	0.333	0.506	0.492	0.351	0.49	0.439	0.412	0.871	
BD 5	0.586	0.706	0.614	0.549	0.651	0.736	0.575	0.493	1.354

Note. BD – Block Design

Table G2

Standardised Residual Covariances for all Analysed Block Design Trials

Trial	BD 13	B 12	BD 11	BD 10	BD 9	BD 8	BD 7	BD 6	BD 5
BD 13	0								
BD 12	0.146	0							
BD 11	0.251	-0.17	0						
BD 10	-0.224	0.025	0.036	0					
BD 9	-0.159	-0.021	0.405	1.013	0				
BD 9	-0.174	0.397	0.075	-0.153	-0.415	0			
BD 7	-0.519	-0.356	-0.379	0.045	0.517	-0.049	0		
BD 6	-0.4	0.128	0.31	-0.586	-0.114	-0.19	0.105	0	
BD 5	0.262	-0.016	-0.266	-0.399	-0.481	0.469	-0.027	0.388	0

Table G3

Factor Score Weights for Final Block Design Model

Latent Factor	BD 13	B _12	BD 11	BD 10	BD 9	BD 8	BD 7	BD 6	BD 5
BLOCK DESIGN C&N	.130	.251	.179	.152	.067	.072	.064	.053	.061
BLOCK DESIGN C&F	.053	.103	.074	.062	.171	.185	.165	.136	.156

Appendix H Revised DCS Criteria and Scoring for Block Design

Table H1

Revised DCS Criteria for Trials 5 – 9 of the Block Design task

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
1.	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established via the provision of a design. The image provides a visual configuration of the design and but required the formulation of sub-goals to achieve various smaller shape designs that comprise the superordinate design
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	The evaluation of the physical state of the colour blocks is required to enable accurate configuration and placement in accordance to the prescribed design
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The physical dimensions of two different solid colour sides and one split colour side of each block must be considered and relationships formed to establish the correct combination of colour sides to achieve the prescribed design
2.	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue.	**2.C**	**The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue**	The operational parameters of the trials change requiring use intermittent exclusive use of split-colour blocks to establish the prescribed design during Trial 11
3.	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Task requires selection of 4 correct block colours (e.g. solid colour side, vs. split red & white sides) amongst four blocks containing 16 sides
4.	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are to establish the response and speed requirements for each trial. Rules are minimal and not uniquely exclusive to this task environment alone (e.g. do not rotate the stimulus book at any time)
5.	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly and is completed before proceeding to the next trial
			(F) Familiar		(N) Novel	
6.	Automaticity	**6.F**	**Task requires the application of implicit knowledge over S→R**	6.N	Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours	The task requires manipulation of four blocks by rotating and positioning them to match the prescribed design
7.	Schematic Demands	**7.F**	**Task requires direct application of implicit fundamental S→R schemas to the task**	7.N	Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas	The task requires only the fundamental knowledge that separate shapes when placed together can collectively establish a new global shape
8.	Episodic Demands	**8.F**	**Minimal to no overt retention of current S→R experience is required**	8.N	Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements	The task does not require the retention of pattern or configuration for successful completion of each trial

Note. text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 5 – 9 of the Block Design	Raw performance score

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 3

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 1
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 2 Maximum = 5	**Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 3 Maximum = 3	**Total score for (N) =** 0 Maximum = 6

Total Score for Novelty (F) + (N) 3

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Familiar

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Familiar

Table H2

Revised DCS Criteria for Trials 10 – 13 of Block Design

ID	Demand Factors		Demand Criteria: (S) Simple		Demand Criteria: (C) Complex	Task Specific Details
1.	Abstraction	1.S		**1.C**		
	(T) Temporal	1T.S	Superordinate goal(s) and responses are made explicit, and the sequence of sub-goals is guided by existing environmental cues	**1T.C**	**Superordinate goal(s) may not be explicit, and the sequence of necessary sub-goals is not offered by the immediate environment**	Superordinate goals are explicitly established via the provision of a design. The image provides a visual configuration of the design and but required the formulation of sub-goals to achieve various smaller shape designs that comprise the superordinate design.
	(P) Policy	1P.S	Lower-order policies contain simple imposed rules that link S→R	**1P.C**	**There is a higher order abstract policy that must be evaluated to enable selection of an appropriate response**	The evaluation of the physical state of the colour blocks is required to enable accurate configuration and placement in accordance to the prescribed design.
	(R) Relational	1R.S	Properties of the stimuli are concrete with a minimal number of independent relationships and dimensions considered to select a response	**1R.C**	**Properties of the stimuli may be variable with many independent relationships and dimensions considered to select a response**	The physical dimensions of two different solid colour sides and one split colour side of each block must be considered and relationships formed to establish the correct combination of colour sides to achieve the prescribed design.
2.	Contextual Stability	2.S	The operational parameters of the task(s) condition(s) are stable. Any changes between trials are alerted by an exogenous cue	**2.C**	**The operational parameters of the task change over the course of administration Changes may not be alerted via an exogenous cue**	The operational parameters of the trials change requiring use of an addition 5 blocks and the exclusive use of split-colour blocks to establish the prescribed design during Trial 11.
3.	Action Rules	3.S	A minimal subset of possible response options are available	**3.C**	**Many alternative response options are available**	Task requires selection of 9 correct block colours (e.g. solid colour side, vs. split red & white sides) amongst four blocks containing 36 sides.
4.	Instructions and rules	**4.S**	**Instructions are given with minimal to no rules and/or a rule set that is not uniquely exclusive to the task environment**	4.C	Instructions and rules may be multiple in quantity or change over the duration of the task and may be uniquely exclusive to the task environment	Instructions are to establish the response and speed requirements for each trial. Rules are minimal and not uniquely exclusive to this task environment alone (e.g. do not rotate the stimulus book at any time).
5.	Dual Nature	**5.S**	**Task requires completion of a singular task**	5.C	Task encompasses a dual or multi-task requirement	Each trial within the task is displayed singularly, and is completed before proceeding to the next trial.
			(F) Familiar		(N) Novel	
6.	Automaticity	6.F	Task requires the application of implicit knowledge over S→R	**6.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge into a series of new explicit task specific behaviours**	The task requires manipulation of nine blocks by rotating and positioning them to match the prescribed design.
7.	Schematic Demands	7.F	Task requires direct application of implicit fundamental S→R schemas to the task	**7.N**	**Task requires the adaptation and/or integration of implicit S→R knowledge to create new task specific schemas**	The task requires the adaptation of fundamental knowledge on shapes to establish a complex design that includes various alternative and novelty shapes.
8.	Episodic Demands	8.F	Minimal to no overt retention of current S→R experience is required	**8.N**	**Task requires the transfer of practiced S→R representations to successfully learn new S→R requirements**	The task requires experience with Trial 12 to be carried to Trial 13 Successful transfer of this S→R from Trial 12 as a diamond configuration will improve application of the diamond configuration to the ambiguity of Trial 13.

Note: text in bold typeface denotes the criteria selection for the task under appraisal.

THE DEMAND CLASSIFICATION (DCS) RECORD SHEET

Name of Test	Scoring Variable
Trials 10-13 of the Block Design	Raw Performance Score

Step 1 **Count the total quantity of COMPLEX Demand Criteria recorded for (T), (P) and (R) and convert to an Abstraction Score**

Total quantity of Complex (T) (P) and (R) Classifications: 3

RAW Score	Abstraction Score
0	1
1	1
2	2
3	2

Abstraction Score 2

Step 2 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(1)	Abstraction Score	(1.S) Simple 1	(1.C) Complex 2
(2)	Contextual Stability Score	(2.S) Simple 1	(2.C) Complex 2
(3)	Action Rules Score	(3.S) Simple 1	(3.C) Complex 2
(4)	Instructions and rules Score	(4.S) Simple 1	(4.C) Complex 2
(5)	Dual Nature Score	(5.S) Simple 1	(5.C) Complex 2
Total Score for each Demand Criteria		**Total score for (S) =** 2 Maximum = 5	**Total score for (C) =** 6 Maximum = 10

Total Score for Complexity (S) + (C) 8

Total Score	Global Complexity Demand
5	Simple Global Demand
6	Simple Global Demand
7	Simple Global Demand
8	Complex Global Demand
9	Complex Global Demand
10	Complex Global Demand

Global Complexity Demand Complex

Step 3 **Circle the score for each Global Demand Criteria**

ID	Demand Factors	Demand Criteria Scores	
(6)	Automaticity Score	(6.F) Familiar 1	(6.N) Novel 2
(7)	Schematic Demands Score	(7.F) Familiar 1	(7.N) Novel 2
(8)	Episodic Demands Score	(8.F) Familiar 1	(8.N) Novel 2
Total Score for each Demand Criteria		**Total score for (F) =** 0 Maximum = 3	**Total score for (N) =** 6 Maximum = 6

Total Score for Novelty (F) + (N) 6

Total Score	Global Novelty Demand
3	Familiar Global Demand
4	Familiar Global Demand
5	Novel Global Demand
6	Novel Global Demand

Global Novelty Demand Novel

GLOBAL DEMAND CLASSIFICATION **Global Complexity Demand** + **Global Novelty Demand** = Complex & Novel

Appendix I Matrices from Confirmatory Factor Analyses of FAS Test

Table J1

Sample Covariances from CFA of FAS Test

	'S' 0-15sec	'S' 16-60sec	A' 16-60sec	'F' 16-60sec	'A' 0-15sec	'F' 0-15sec
'S' 0-15sec	3.218					
'S' 15-60sec	.552	1.427				
A' 15- 60sec	.227	.670	1.107			
'F' 15-60sec	.518	.620	.581	1.243		
'A' 0-15sec	1.128	.636	.608	.848	2.832	
'F' 0-15sec	.956	.901	.874	.640	1.493	4.791

Table J2

Standardised Residuals from CFA of FAS Test

	'S' 0-15sec	'S' 16-60sec	A' 16-60sec	'F' 16-60sec	'A' 0-15sec	'F' 0-15sec
'S' 0-15sec	0.000					
'S' 15-60sec	0.213	0.000				
A' 15-60sec	-1.213	0.253	0.000			
'F' 15-60sec	0.235	0.203	-0.072	0.000		
'A' 0-15sec	0.368	-0.525	-0.380	0.770	0.000	
'F' 0-15sec	-0.236	0.464	0.704	-0.327	-0.154	0.000

Table J3

Factor Score Weights from CFA of FAS Test

Latent Factor	'S' 0-15sec	'S' 16-60sec	A' 16-60sec	'F' 16-60sec	'A' 0-15sec	'F' 0-15sec
FAS S&N	.028	.252	.317	.269	.078	.034
FAS S&F	.092	.098	.123	.104	.257	.112

Note. S&N = Simple and Novel; S&F = Simple and Familiar.

Appendix J Matrices from Confirmatory Factor Analyses of Austin Maze.

Table I1

Squared Multiple Correlations of M1

Trial	Estimate
Trial 6	.751
Trial 5	.818
Trial 9	.722
Trial 8	.789
Trial 7	.789
Trial 4	.807
Trial 3	.693
Trial 2	.325

Table I2

Sample covariance for Austin Maze model M0

Trial	Trial 5	Trial 9	Trial 8	Trial 7	Trial 4	Trial 3
Trial 5	11.948					
Trial 9	6.730	9.342				
Trial 8	7.626	7.268	9.342			
Trial 7	9.249	8.765	8.453	12.883		
Trial 4	10.017	6.978	7.542	9.267	12.831	
Trial 3	11.212	7.471	8.193	10.019	11.712	18.863

Table I3

Standardised Residuals for Austin Maze model M0

Trial	Trial 5	Trial 9	Trial 8	Trial 7	Trial 4	Trial 3
Trial 5	.000					
Trial 9	-.376	.000				
Trial 8	.296	.161	.000			
Trial 7	.255	.080	-.183	.000		
Trial 4	-.047	-.309	.088	.121	.000	
Trial 3	-.010	-.481	-.046	-.045	.110	.000

Table I4

Factor Score Weights for Austin Maze

Latent Factor	Trial 5	Trial 9	Trial 8	Trial 7	Trial 4	Trial 3
AUSTIN MAZE C&F	.024	.081	.085	.103	.021	.010
AUSTIN MAZE C&N	.106	.019	.020	.024	.091	.042

Note. C&F= Complex and Familiar; C&N = Complex and Novel.

Appendix K Re-scaled Factor Score Weights for Latent Factors of Study 1b

Table K1

Original and Rescaled Factor Weights for Block Design

Latent Factor	Factor Score Weights									
	BD 13	BD 12	BD 11	BD 10	BD 9	BD 8	BD 7	BD 6	BD 5	Total
BLOCK DESIGN C&N										
Original weights	.130	.251	.179	.152	.067	.072	.064	.053	.061	1.029
Proportioned weights	.126	.244	.174	.148	.065	.070	.062	.052	.059	1.000
BLOCK DESIGN C&F										
Original weights	.053	.103	.074	.062	.171	.185	.165	.136	.156	1.105
Proportioned weights	.048	.093	.067	.056	.155	.167	.149	.123	.141	1.000

Note. Proportioned Weights totalled = 1. BD = Block Design; C&N = Complex and Novel, C&F = Complex and Familiar.

Table K2

Original and Rescaled Factor Weights for FAS Test

	Factor Score Weights						
Latent Factor	‘S’ 0-15sec	‘S’ 15-60sec	A’ 15- 60sec	‘F’ 15-60sec	‘A’ 0-15sec	‘F’ 0-15sec	‘S’ 0-15sec
FAS SIMPLE & NOVEL							
Original weights	.028	.252	.317	.269	.078	.034	.978
Proportioned weights	.029	.258	.323	.275	.080	.035	1.00
FAS SIMPLE & FAMILIAR							
Original weights	.092	.098	.123	.104	.257	.112	.786
Proportioned weights	.117	.125	.156	.132	.328	.142	1.00

Note. Proportioned Weights totalled = 1; S&N = Simple and Novel; S&F = Simple and Familiar.

Table K3

Original and Rescaled Factor Weights for Tower of Hanoi

	Factor Score Weights					
Latent Factor	Trial 11	Trial 8	Trial 7	Trial 5	Trial 3	Total
TOH COMPLEX & FAMILIAR						
Original weights	0.009	0.125	0.046	0.07	0.047	0.297
Proportioned weights	.030	.421	.155	.236	.158	1.000

Note. Proportioned Weights totalled = 1. TOH = Tower of Hanoi.

Table K4

Original and Rescaled Factor Weights for Austin Maze

	Factor Sore Weights						
Latent Factor	Trial 5	Trial 9	Trial 8	Trial 7	Trial 4	Trial 3	Total
AUSTIN MAZE C&N							
Original weights	.106	.019	.020	.024	.091	.042	0.302
Proportioned weights	.352	.063	.066	.079	.301	.139	1.000
AUSTIN MAZE C&F							
Original weights	.024	.081	.085	.103	.021	.010	.324
Proportioned weights	.074	.250	.262	.318	.065	.031	1.00

Note. Proportioned Weights totalled = 1. C&N = Complex and Novel, C&F = Complex and Familiar.

Appendix L Covariance Matrices and Standardised Residuals of S&F Model

Table L1

Sample Covariance Matrix of S&F model

Variable	Map Search 0-60secs	Test of d2	5-point 0-60secs	FAS S&F	TMT A	Stroop Test – Words
Map Search 0-60secs	132.261					
Test of d2	76.286	1353.132				
5-point 0-60secs	2.692	47.730	22.394			
FAS S&F	-0.115	9.428	0.620	1.167		
TMT-A	14.719	64.466	4.165	1.263	42.967	
Stroop Test – Words	3.485	37.701	3.489	0.857	5.441	5.779

Note: FAS S&F = FAS Test Simple and Familiar Score.

Table L2

Standardised Residual Covariances of S&F model

Variable	Map Search 0-60secs	Test of D2	5-point 0-60secs	FAS S&F	TMT A	Stroop Test – Words
Map Search 0-60secs	.000					
Test of D2	.571	.000				
5-point 0-60secs	-.333	.382	.000			
FAS S&F	-.943	-.033	-.396	.000		
TMT A	1.007	-.037	-.468	-.071	.000	
Stroop Test – Words	-.290	-.148	.105	.266	.041	.000

Note. FAS S&F = FAS Test Simple and Familiar Score.

Appendix M Covariance Matrices and Standardised Residuals of S&N Model

Table M1

Sample Covariance Matrix of S&N model

Variable	Digit Span-Backwards	FAS S&N	5-point 61-120secs	Visual Span-Backwards	Visual Elevator
Digit Span-Backwards	5.663				
FAS S&N	.523	.814			
5-point 61-120secs	.190	.413	9.625		
Visual Span -Backwards	.901	.134	.465	3.063	
Visual Elevator	.527	.263	.612	.362	.769

Note. FAS S&N = FAS Test S&N Score

Table M2

Standardised Residual Covariances of S&N model

Variable	Digit Span-Backwards	FAS S&N	5-point 61-120secs	Visual Span-Backwards	Visual Elevator
Digit Span Backwards	.000				
FAS S&N	.484	.000			
5-point 61-120secs	-.892	.109	.000		
Visual Span Backwards	.823	-.749	-.084	.000	
Visual Elevator	-.252	.007	.297	.074	.000

Note. FAS S&N = FAS Test Simple and Novel Score.

Appendix N Covariance Matrices and Standardised Residuals of C&F Model

Table N1

Sample Covariance Matrix of C&F model

Variable	BD C&F	TOH	Austin Maze C&F	TMT-B	Map Search 61-120secs
BD C&F	2.507				
TOH	.634	7.070			
Austin Maze C&F	1.845	1.108	8.683		
TMT-B	6.548	9.663	5.084	247.340	
Map Search 61-120secs	3.787	3.879	4.362	39.478	70.363

Note. BD = Block Design; TOH = Tower of Hanoi; TMT =Trail Making Test; C&F = Complex and Familiar.

Table N2

Standardised residual Covariances of C&F model

Variable	BD C&F	TOH	Austin Maze C&F	TMT B	Map Search 61-120secs
BD C&F	.000				
TOH	-.534	.000			
Austin Maze C&F	.691	-.084	.000		
TMT-B	-.229	.978	-.984	.000	
Map Search 61-120secs	-.241	.298	-.495	.965	.000

Note. BD = Block Design; TOH = Tower of Hanoi; TMT = Trail Making Test; C&F = Complex and Familiar.

Appendix O Covariance Matrices and Standardised Residuals of C&N Model

Table O1

Sample Covariance Matrix of C&N Model

Variable	Stroop Test: Colour-Word	ECR	BD C&N	Austin Maze C&N
Stroop Test: Colour-Word	25.582			
ECR	4.341	4.940		
BD C&N	2.722	1.711	3.261	
Austin Maze C&N	.722	2.616	2.292	10.760

Note. ECR = Elevator Counting Reversal; BD= Block Design; C&N = Complex and Novel.

Table O2

Standardised residual Covariances of C&N model

Variable	Stroop Test: Colour-Word	ECR	BD C&N	Austin Maze C&N
Stroop Test: Colour-Word	.000			
ECR	.264	.000		
BD C&N	-.322	.000	.000	
Austin Maze C&N	.000	-.242	.296	.000

Note. ECR = Elevator Counting Reversal; BD= Block Design, C&N = Complex and Novel.

Appendix P Re-scaled Factor Score Weights of GDC Factors from Study 2

Table P1

Original and Rescaled Factor Weights for S&F Model

	Factor Score Weights					
Latent Factor	Test of D2	5-point 0-60secs	FAS S&F	TMT A	Stroop Test Words	Total
SIMPLE & FAMILIAR						
Original weights	.007	.028	.131	.024	.215	.405
Proportioned weights	.017	.069	.323	.059	.532	1.00

Note. Proportioned Weights totalled = 1. FAS S&F = FAS Test Simple and Familiar Score.

Table P2

Original and Rescaled Factor Weights for S&N Model

	Factor Score Weights					
Latent Factor	Digit Span-Backwards	FAS S&N	5-point 61-120secs	Visual Span-Backwards	Visual Elevator	Total
SIMPLE & NOVEL						
Original weights	0.078	0.266	0.038	0.082	0.572	1.036
Proportioned weights	.075	.257	.037	.079	.552	1.00

Note. Proportioned Weights totalled = 1. FAS S&N = FAS Test Simple and Novel Score; S&N = Simple and Novel.

Table P3

Original and Rescaled Factor Weights for C&F Model

	Factor Score Weights					
Latent Factor	BD C&F	TOH	AM C&F	TMT-B	Map Search 61-120secs	Total
COMPLEX & FAMILIAR						
Original weights	0.277	0.047	0.08	0.012	0.026	0.442
Proportioned weights	.627	.106	.181	.027	.059	1.00

Note. Proportioned Weights totalled = 1. BD =Block Design; TOH = Tower of Hanoi; TMT – Trail Making Test; AM= Austin Maze. C&F = Complex and Familiar.

Table P4

Original and Rescaled Factor Weights for C&F Model

	Factor Score Weights				
Latent Factor	Stroop Test Colour-Word	ECR	BD C&N	AM C&N	Total
Complex & Novel					
Original weights	0.061	0.147	0.149	0.101	0.458
Proportioned weights	.133	.321	.325	.221	1.00

Note. Proportioned Weights totalled = 1. BD = Block Design; ECR = Elevator Counting Reversal; BD= Block Design; AM= Austin Maze; C&N = Complex and Novel.

Appendix Q Covariance Matrices and Standardised Residuals of SEM Model

Table Q1

Sample Covariances for Full structural model analyses

	S&N	S&F	C&N	C&F
S&N	1.532			
S&F	1.891	4.445		
C&N	.981	1.782	2.507	
C&F	.781	1.437	2.186	2.991

Table Q2

Standardised residual Covariances for Full structural model analyses

	S&N	S&F	C&N	C&F
S&N	-.094			
S&F	.026	.000		
C&N	.133	-.063	.000	
C&F	.072	-.035	.000	.000