THE NEURAL UNDERPINNNGS OF MENTAL ATTENTION CAPACITY

MYLANN GUEVARA

A THESIS SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO
OCTOBER 2018

Abstract

Working memory refers to the online maintenance and manipulation of information to solve problems; a function that is constrained by a limited mental attention capacity (M-capacity). The Theory of Constructive Operators (TCO) describes M-capacity as an operator that can boost task relevant information in the brain while task irrelevant information is inhibited in order to facilitate successful task performance. The current investigation aimed to study the neural underpinnings of M-capacity in the brain, specifically looking at how neural activation is modulated as a function of cognitive load and the domain of stimuli; with regard for how this relationship is affected by critical factors predicted by the TCO. A dynamic relationship between brain activity and cognitive load was found, where different neural networks became engaged and disengaged depending on cognitive load, and this relationship interacted with the domain of stimuli in the bilateral fusiform gyri and right middle occipital gyrus. These observations justify a number of future investigations to further understand how the modulation of neural networks in the brain relates to M-capacity generally, and individual differences in M-capacity.

Acknowledgements

I would like to firstly acknowledge my parents who have supported me throughout my entire life and always instilled in me the idea that I can do anything that I put my mind to. I would like to specifically thank my dad for always reinforcing my interest in science from the time that I was very young. Thank you to my sister Breanne Guevara who has inspired me to believe that self-improvement is possible and to always make me laugh on Instagram. Thank you to my life partner Jahred Weekes who supports me every single day in so many ways. I would also like to acknowledge my grandparents and aunt who have played a huge role in supporting me throughout my studies and have always been there for me when I needed help.

I would like to thank my supervisor Dr. W. Dale Stevens who has guided me through the immensely difficult task of designing an MRI experiment for the first time. I have learned so much from my supervisor and I am so grateful for his guidance and never ending support in my academic endeavours. It doesn’t matter what time of day or night, you have been there to answer my questions and give me transparent feedback and for that I am so lucky. I look up to you as a researcher and scientist and I hope to know as much about MRI as you do some day! I couldn’t have completed this thesis without you!

I would also like to thank my collaborators Dr. Marie Arsalidou and Dr. Juan Pasqual-Leon who have been instrumental in guiding me through completing this study and conducting research within Dr. Pasqual-Leon’s theoretical framework. I am extremely privileged to be able to work with two heavyweights in the field of developmental psychology and an academic descendant of Piaget himself. You both have believed in me from the start and for that I am extremely grateful. I hope I will make you both proud!

I would like to thank the MRI Technologist at the York University MRI research Facility Joy Williams who helped me collect all of my MRI data and assisted Dr. Stevens and myself in choosing optimal MRI protocol.

I would also like to thank all of the Undergraduate student volunteers including Adina Levi who helped me recruit and test such a large number of participants in a short period of time. I could not have finished my study before the scanner upgrade without you!

Finally, I would like to thank my lab mates Lily, Naail, and Katherine who have seen me stumble and get back up for this last year of my Master’s. You all have been a huge influence on my life and without your moral support along the way, I am not sure that I would have my sanity today.

Table of Contents

List of Tables

List of Figures

List of Appendices

Introduction

Working memory (WM) refers to the ability to maintain relevant information online so that it can be manipulated and used to solve problems (Engle, 2001). Mental attention capacity (M- capacity) refers to the limited amount of task-relevant information that can be maintained, despite the presence of task-irrelevant distractions (Pascual-Leone, 1970; Pascual-Leone & Johshon, 2005). Since the mid-20th century, scientists have investigated the nature of WM and M-capacity in ways characterized by the technology available at the time. With the evolution of research methods followed the specification of research questions, allowing us to inform our previously standing theoretical models of WM and the nature of M-capacity. The current project aims to investigate the neural underpinnings of M-capacity: 1) within a theoretical framework that was developed before the era of functional magnetic resonance imaging (fMRI), and 2) using an experimental paradigm that has previously been used with children. This investigation will therefore help to extend our knowledge of M-capacity across young adults and previous research with children in a seamless way while also informing the theoretical framework at hand.

The Theory of Constructive Operators. Dr. Juan Pascual-Leone has developed the Theory of Constructive Operators (TCO), which posits that the maturation of M-capacity, in part, causes the development of intelligence (Pascual-Leone & Johnson, 2005). Fundamentally, the capacity to handle a greater number of task-related elements is critical for being able to solve more complex and difficult problems (Pennings & Hessels, 1996). According to the TCO, operators are generic functions that can be applied to content bearing schemes (which are cognitive structures of information about the world that help individuals understand their perceptions), in order to manipulate the contents of WM (Pascual-Leone & Johnson, 2005). Within the TCO framework, M-capacity is an M-operator that boosts the activation of task-relevant schemes, while the I-operator can supress irrelevant schemes (Pascual-Leone & Johnson, 2005). Successful WM function in the context of the TCO involves the simultaneous forces of the “I” and “M” operators, supressing and activating the correct schemes, so that relevant information may be accessible and sources of interference may be nullified.

The Emergence of WM. Before the mid-20th century, memory was considered to be one unitary function, but when neurological patients displayed selective memory deficits, the support for a dichotomized model of the memory system prevailed (Eriksson et al., 2015). Attkinson and Shiffrin depicted the brain as a computer, and the flow of information linearly, from sensory memory into short-term memory, and eventually into long-term memory, by way of active mental processes (Atkinson & Shiriffin, 1968). Evidence from frontal lobe patients supported the construct of WM (Owen et al., 1990) as a differentiation from the basic construct of short-term memory/retention (STM) (Kirchner, 1958; Crowder, 1982). WM necessarily involves the mental selection (top-down attention) and manipulation of incoming information; whereas STM involves short-term storage without the requirement of mental operation (Crowder, 1982). As neuroimaging techniques advanced, evidence that cognitive functions rely on distributed processing in the brain accumulated, because the ability to observe and quantify interactions among brain regions became increasingly more precise in terms of temporal and spatial resolution (Gazzaley et al., 2004). Consequently, cognitive neuroscience investigations that are conducted today more commonly highlight how unitary calculations come together to accomplish distributed processing for complex functions (Gazzaley et al., 2004; Fox et al., 2005).

As a complex function, it is widely accepted that WM employs a variety of different processes and mental representations, which come together in the service of some goal (Fuster, 2009; Gazzaley et al., 2004). Critically, it is the very nature of the information to be manipulated and the intended goal for that information that will determine the selection of mental processes and the ensemble of brain regions that become active (Eriksson et al., 2015). Posterior association areas that support memory for modality specific information are involved in WM tasks within the same modalities (Eriksson et al., 2015). Long-term memory (LTM) for visuospatial information involves activation of the superior parietal lobe and WM tasks involving visuospatial processing activate this area also (Eriksson et al., 2015). Gazzaley and colleagues (2004) found evidence to suggest that activity of the dorsal and ventrolateral prefrontal cortex (PFC), intraparietal sulcus (IPS), thalamus, hippocampus, and other occipitotemporal regions were all correlated with fusiform gyrus activity, and involved in the active maintenance of information during a WM delay period. WM involves an interaction of top-down selective attention processes and the temporary activation of LTM representations (Awh et al., 2006; Eriksson et al., 2015). Brain regions within the PFC and parietal lobe are thought to exert this top-down attentional control (D'Esposito & Postle, 1999; Awh et al., 2006) to facilitate the suppression of irrelevant information in posterior sensory areas (Ungerleider, 2000). The distributed processing view of WM (Fuster, 2009) is consistent with the TCO, which holds that

WM involves the interplay between operators and schemes. Here, the “M” and “I” operators are comparable to top-down attentional processes which act to temporarily boost/activate and suppress schemes, respectively, which fall under the umbrella of LTM representations (Pascual-Leone, 1970).

What is less agreed upon between the TCO and other views of WM is whether or not the capacity of WM is domain specific, considering that different brain areas are expected to serve WM during tasks within different domains. The outcome of this debate holds consequences for how M-capacity should be operationalized because it will mean that M-capacity should be measured in domain specific/non-specific ways and could constrain our ability to compare M-capacity across domains.

The Multiple Component Model. A domain specific view of M-capacity would mean that an individual’s M-capacity for colours could be different from their M-capacity for numbers, for example. This would suggest that the source of limitation for M-capacity is as distributed as its parts and may be localized in brain regions that are specialized for processing particular kinds of stimuli. The Multiple Component Model (Baddeley & Hitch, 1974) was one of the first models to dichotomize WM, where the interaction of domain specific components is orchestrated by a higher order central executive. This model claimed that the capacity limitations of WM would be domain specific, corresponding to the separate components of the model such as the: visuospatial sketchpad and the phonological loop, dedicated to the active maintenance of visuospatial and acoustic information, respectively.

Evidence to support the domain specific view of M-capacity comes from the dual task paradigm where a participant is required to perform two separate functions simultaneously (Cocchin et al., 2002). An example of this paradigm is the word span task, where participants are asked to read a list of sentences, then later recall the last word of each sentence (Cocchin et al., 2002). If similar performance is observed during dual task and individual task completion, then it is believed that separate cognitive resources are serving each domain, explaining the lack of trade-off between accuracy and completing an additional irrelevant task (Cocchin et al., 2002).

An individual has high functioning WM if their accuracy for commonly facilitated functions remains high during dual task completion (Cocchin et al., 2002). An issue with the dual task paradigm is that there is seldom sufficient control of executive demand across the two tasks; the executive demand involved in reading a sentence is not equal to that of remembering a list of single words. In the case of a high-WM individual, this discrepancy in executive demand leaves an alternative explanation for the lack of interference, namely that the tasks are easy enough to be performed in tandem. By this logic, once the sum of executive demand between the two functions exceeds an individual's threshold, behavioural performance would be expected to decline and there is no way to know for sure whether or not the behavioural decline in one individual is caused by the same factor in another individual.

Measuring M-capacity. Recall that according to the TCO, M-capacity refers to the extent to which information can be maintained and manipulated in WM, which simultaneously engages the "I" and "M" operators (Pascual-Leone & Johshon, 2005). M-capacity is therefore most authentically measured under misleading conditions, where the inhibition of misleading (task-irrelevant) schemes and boosting of facilitating (task-relevant) schemes is required (Arsalidou et al., 2010). The TCO defines a facilitating task as one that involves more task-relevant (facilitating) elements, resulting in less interference; whereas a misleading task involves more misleading (task-irrelevant) elements in addition to any facilitating ones (Arsalidou et al., 2010). Specifically, the failure to supress misleading schemes will result in error, a characteristic that is used to draw insight about an individual's ability to execute mental control, exhibited in their behavioural accuracy on a misleading task (Arsalidou et al., 2010).

While the TCO indirectly implies a certain duality, in that it requires that relevant/irrelevant information be boosted/supressed, respectively, the critical difference from a dual task paradigm is that in order for this duality to provide a valid measurement of M-capacity, the executive demand must be held constant (Arsalidou et al., 2010). Executive demand refers to the type of and number of executive processes/operations that need to be engaged to complete a task; while M-demand refers to cognitive load/set size, or the amount of information that must be maintained at once. Arsalidou and colleagues (2010) demonstrated that a misleading task with parametrically varied M-demand and controlled executive demand provided a measurement of M-capacity that was in-line with a previously established measure; whereas a facilitating version of the same task with the same M-demand did not. Arsalidou and colleagues (2010) defined M-demand in the facilitating task as n + 1 where n is the number of task-relevant elements and the "1" refers to the level of executive demand. M-demand in the misleading task was defined as n + 2 where there was a one unit increase in executive demand (Arsalidou et al., 2013). These results suggest that the difference between the misleading and facilitating tasks stemmed from a difference in executive demand (the degree to which the I-operator was engaged) and not cognitive load. Moreover, the decline in accuracy associated with higher cognitive load for participants between the ages of 5-16 years old related to the participants' developmental stage linearly in the misleading task, but not in the facilitating task, suggesting that executive demand does not interact with set size in a linear fashion to affect behavioural performance (Arsalidou et al., 2010). In line with the TCO, facilitating-WM tasks with a corresponding set size will have low executive demand because the application of the I-operator is not required. This characteristic means that a misleading task will have more construct validity than a facilitating task when quantifying M-capacity (Arsalidou et al., 2010).

N-back Paradigm. A common paradigm used to measure WM is the n-back paradigm (Kirchner, 1958), where participants are required to make a same/different judgement about the current stimulus (probe) relative to a target stimulus that appeared some number (n) of stimulus presentations prior (Owen et al., 2005). N-back tasks can vary cognitive load in different ways, either by varying n (the number of serial stimulus presentations between the target and the probe) (Callicott et al., 1999), or by varying set size (the number of elements within a stimulus that must be compared to make a correct judgement) (Arsalidou, 2010). The former option will vary difficulty by increasing executive demand because additional mental processes like updating and rehearsing visual stimuli are required in order to compare stimuli that are two and three stimulus presentations apart (Arsalidou et al., 2010). The varying executive demand will be expected to interact with the set size of the task, convolving the measurement of M-capacity with executive control (Arsalidou et al., 2010). To measure M-capacity within the TCO framework, the latter option has been implemented because it holds executive demand constant, allowing for more direct measurement of M-capacity. Furthermore, the manipulation of set size allows for a greater range of difficulty levels; which is ideal for capturing individual differences within and across populations, providing insight into the maturation and manifestation of this cognitive resource.

Previous research investigating M-capacity has demonstrated that it increases across human development from the ages of 3-16 years old, with an average one-unit increase every two years, until an eventual average of seven-units (Pascual-Leone & Johnson, 2005; Arsalidou et al., 2010). Importantly, this linear trajectory of M-capacity growth is in accordance with Piaget's stages of development, and has been presented as a quantitative explanation for the qualitative transition from one developmental stage to the next (Pennings & Hessels, 1996). One benefit of the current research is to apply this same parametric approach to investigate the neural underpinnings of M-capacity in young adults so as to inform future investigation in other populations.

Theoretical M-capacity and Functional Neuroimaging of Neural Networks

Before the prevalence of functional neuroimaging, most of what we knew about the organization of function in the human brain came from lesion patients (Scoville & Milner, 1957), and primate research (Funahashi et al., 1998). Today, functional neuroimaging allowing for whole brain acquisition indexed by dynamically shifting cognitive states has uncovered certain key groups of brain areas that go into and out of synchrony according to the cognitive orientation of the brain (Shulman et al., 1997; Corbetta & Shulman, 2002). These groups of brain areas, known as neural networks, are believed to underpin higher order cognitive functioning by mobilizing ensembles of brain areas to carry out complex functions in a distributed manner. Hence, a functional neural network is simply a group of brain regions whose activation amplitude fluctuations covary more with each other than with other groups of covarying brain regions (Fox et al., 2005; Spreng et al., 2016). In order to understand a complex function like WM, the consideration of relevant neural networks is imperative, given the inherent multidimensionality in the operational definition of WM.

Until recent years, the TCO has been grounded in behavioural observations (Arsalidou et al, 2010). This changed when Arsalidou and colleagues (2013) used fMRI to investigate the neural responses to parametrically increased M-demand during the same misleading task that they studied before (Arsalidou et al., 2010). As difficulty increased, parametric increases in activation were observed in the bilateral fusiform gyrus, and a number of brain regions within the frontoparietal control network: middle PFC, cingulate gyrus, left precentral gyrus, and precuneus. These difficulty-related increases in activation were accompanied by parametric decreases in default network brain regions, which included the medial PFC, left posterior cingulate, right superior temporal gyrus, and the bilateral medial frontal gyrus. Given that the default network is reliably supressed during externally focussed cognition (Raichle et al., 2010), and activated during internally focussed cognition (Spreng et al., 2010), the researchers hypothesized that parametric suppression of default network brain areas reflected a proportional increase in task-driven recruitment of M-capacity (Arsalidou et al., 2013). Taken together, these observations are in line with other evidence suggesting that the frontoparietal control network exerts top-down inhibitory control on the default network to focus externally driven attention (Spreng et al., 2010).

Since the 2013 investigation of M-capacity, our understanding of neural networks and M-capacity has expanded. Van Snellenberg and colleagues (2015) used a self-ordered WM task (SOT) that presented participants with an array of eight line drawings at once; requiring them to iteratively select one line drawing that they have not previously selected, with the images being scrambled after each choice. Cognitive load increases as the participant needs to remember more of the previous images that they have selected despite the inconsistency of their spatial location. The goal was to investigate the neural correlates of exceeding one's WM-capacity.

Van Snellenberg and colleagues (2015) found that cognitive load did not modulate blood-oxygen-level dependant (BOLD) signal monotonically, but rather, dynamic changes in different neural networks occurred in response to parametrically increasing cognitive load to the point of capacity. Notably, brain regions associated with the default network showed a U-shape function, where suppression of the default network increased as a function of increasing cognitive load until WM reached capacity limitations, at which point the network increased in activation. Another network referred to as the WM network, (consisting of the dorsolateral PFC, posterior parietal cortex, and pre-supplementary motor area), showed an inverted U-shape function in the opposite direction, where BOLD signal increased with cognitive load until near capacity limitations of six to seven steps into the SOT, at which point BOLD signal decreased. Of significant importance is the fact that the researchers only included accurate trials in their analysis, thereby discounting the alternative explanation that participants disengaged from the task when it became too difficult.

Superficially, the findings from Van Snellenberg and colleagues (2015) appear to be at odds with the TCO and Arsalidou and colleagues (2013). However, if it is accepted that executive demand remained constant in the SOT across parametrically increased cognitive load, then deviance from the Arsalidou and colleagues (2013) findings may be due to a difference in the methods or a difference in cognitive load. Arsalidou and Colleagues (2013) assessed cognitive load from one-to-six elements, which corresponds to three-to-eight capacity units, as defined by the TCO (Arsalidou et al., 2010), whereas Van Snellenberg and colleagues (2015) assessed cognitive load from one to eight elements; but it is not evident how this range corresponds to capacity units as defined by the TCO. Therefore, it is possible that the linear relationship between BOLD signal and cognitive load found in the investigation by Arsalidou and Colleagues (2013) could be the early stage of a curvilinear function that might be uncovered by 1) increasing statistical power or 2) expanding the range of cognitive load assessed in the n-back paradigm used to test M-capacity.

Within the TCO framework, an individual's M-capacity across development increases to an average of seven units, which means that the n-back paradigm used by Arsalidou and Colleagues (2013), which tested a maximum M-capacity of eight units, possessed the range in cognitive load to test average participants at their maximum M-capacity. Furthermore, the investigation by Arsalidou and Colleagues (2013) included a total of ten participants, whereas the investigation by Van Snellenberg and colleagues (2015) included thirty-six participants. Therefore, it is more plausible that an increase in statistical power (an increase in the ability to detect a true experimental effect) using the same n-back paradigm as Arsalidou and colleagues (2013) might lead to observations that are in line with Van Snellenberg and colleagues (2015).

Arsalidou and colleagues modeled the BOLD signal in response to parametrically increasing cognitive load linearly, such that voxels exhibiting a linear relationship with cognitive load were identified in a univariate linear contrast. BOLD signal was then extracted from voxels at the locations of peak activation across the six levels of cognitive load, depicting the linear relationship between BOLD signal and task difficulty. This approach omits voxels which did not show a linear relationship with cognitive load. Therefore, the way in which the expected BOLD response is modeled in a WM task will characterize the conclusions that can be drawn.

The Current Study

Objective. The present study is an investigation of the neural underpinnings of M-capacity. The data collected during this study afford the possibility of implementing many different analyses, a subset of which is the aim of the current study. The current investigation aims to: (1) optimize the previously used paradigm by Arsalidou and colleagues (2010) for fMRI; (2) increase the sample size and update or replicate previous fMRI results by Arsalidou and colleagues (2013); (3) compare neural responses across both facilitating and misleading tasks, and (4) an additional misleading task with a new domain of stimuli. These investigations will further our understanding of whether or not M-capacity is domain specific in the brain; and whether or not M-demand relates to brain activation in a monotonic linear fashion, and if so, where and how.

Hypotheses. Given the extensive optimization of the n-back paradigm in the current study, relative to that used by Arsalidou and colleagues (2010) for fMRI, we expect to: (1) observe similar linear increases and decreases in frontoparietal and default network regions, respectively, keeping in mind that it is not clear how the modulation of BOLD signal will change with an increase in power; (2) we also expect to see an interaction between difficulty-level and task-type, even when there are the same number of elements to attend to, illustrating the additional engagement of the (I)-operator in a misleading task; and (3) It is still unclear whether or not the new-domain misleading task will yield a different M-capacity measure from the other previously established misleading task; but if the measurement of M-capacity remains the same, this will support the view of the TCO that M-capacity is not domain-specific.

Methods

Participants

Thirty healthy adults (16 female, 14 male) between the ages of 20-30 (Mean age= 23, *SD*= 2.85) were recruited through posters and word of mouth. Ethics were approved for this project by the York University ethics board and informed consent was obtained before participation. Participants were compensated $20.00 for participating in the study. Exclusionary criteria for this study were: diagnosed neurological disorders, brain injury, left-handedness, abnormal vision that could not be corrected with basic lenses, non-proficiency with the English language, any conditions that made the participant incompatible with the MRI environment (such as a pacemaker or claustrophobia), and colour blindness (tested during the instructional phase).

Materials and apparatus

MRI simulator. The MRI simulator is an apparatus that is designed to emulate the real MRI environment. The simulator has the same dimensions as the bore of the MRI scanner and plays audio recordings of MRI-scans through speakers embedded in the walls of the bore. The bed slides in and out of the bore by button press to simulate the feeling of going into the MRI-scanner. A mock head coil is placed over the participant’s head, and a mirror attached to it allows them to see the screen that is behind their head while lying in the simulator, as per a typical MRI setup. A projector, which rests on a mounted stand, projects a visual display onto the screen behind the participant’s head and a button box that is a replica of the one used in the MRI-scanner is used to collect responses in the simulator. The simulator is also equipped with a motion tracking apparatus that interfaces with software to provide the participant with real-time feedback of their own head motion via the visual display.

Participants completed a practice version of the experimental task in the MRI simulator to habituate them to the MRI environment and reduce the potential effects of novelty and/or anxiety on neural activity during subsequent testing in the actual MRI environment. Participants also completed motion reduction training in the MRI simulator, which requires participants to execute instructed movements systematically, while monitoring their own head motion; the objective is to convey to the participant the importance of not moving while in the MRI environment.

Experimental tasks

Arsalidou and colleagues (2010; 2013) have developed the Colour Matching Task (CMT) to measure M-capacity in children and adults. They have since also developed the number matching task (NMT), which uses the same paradigm as the CMT, but with numbers being the elements to be attended, instead of colours. These are adapted 1-back tasks (present a series of stimuli and require the participant to actively make same/different judgements in real- time), which parametrically increase M-demand by increasing the number of elements that must be attended to at once.

Colour Matching Task (CMT). The CMT presented participants with a sequence of stimuli and required them to make a same/different judgement by indicating whether or not the current stimulus had the same set of relevant colours as the previous one (Figures 1 and 3). Participants were required to ignore the colours green and blue while attending to the colours of interest: brown, grey, orange, pink, purple, red, and yellow. A stimulus was considered "same" if it had elements that consisted of all the same colours of interest appearing in the prior stimulus, irrespective of the location and number of elements of each colour. This means that even if blue elements in a prior stimulus change to the colour green in the current stimulus, so long as all of, and only, the colours of interest are still present, the stimulus is considered to be "same". There are a total of six difficulty levels that index M-demand, determined by the number of colours of interest that need to be attended to at once. There are facilitating and misleading versions of the CMT called the CMT balloon and the CMT clown, where the stimuli comprised balloon and clown pictures, respectively (Figure 1). The CMT balloon is a "facilitating task" because the balloons are homogenous in shape and size and their only distinguishing feature is their colour.

The M-demand for this task is modelled as n + 1, where n is the number of colours of interest and "1"refers to the level of executive demand. The CMT clown is a "misleading task", because the size and shape of the coloured elements are variable. Furthermore, participants were instructed to also ignore the face of the clown, which provides an additional element that must be supressed by the I-operator. The M-demand of this task is modelled as n + 2 where the misleading nature of this task results in an executive demand level of 2, one unit higher than the CMT balloon.

Participants made their responses on a right-handed button box with four buttons. Only three of the buttons were active and they were instructed to press the first button with their index finger to indicate that a stimulus is the "same", and the second button with their middle finger to indicate that a stimulus is "different". Participants were instructed to press the third button with their ring finger when they were unable to make a comparison with the current stimulus, which occurred when there was no prior stimulus (the first stimulus in a series), or when there was a control stimulus (entirely comprising elements that are blue and green). The purpose of the third button was to match brain activation associated with button-presses across all trial-types.

The CMT was a 1-back block design with four runs. Each run had seven blocks presented in random order. One of those blocks was a control block presenting only control stimuli. The other six blocks were assigned to each of the six difficulty levels. There was one block of each difficulty level and one control block in each run, thus, four iterations of each block-type across the entire task.

Each block presented the participant with a total of 8 stimuli (56 stimuli per run, 224 stimuli per task) where each stimulus was presented for 3 s with an inter-stimulus interval of 1 s (Figure 3). Task blocks had seven stimulus presentations of interest (excluding the first stimulus presentation), where the participant made a same/different judgement (42 stimulus presentations of interest per run, 168 presentations of interest per task, 28 presentations of interest per difficulty level. Each block concluded with a 1.5 s block-offset cue (black + on a white background) that indicated the end of the block. Beginning each run and interleaving the blocks was a 10 s period of fixation (black X on a white background) followed by a 2.5 s block-onset cue (black + on a white background). Therefore, each block spanned 32 s and each run spanned 5 min and 24 s; a total of 4 runs (one whole task) spanned 21 min and 36 s.

Number Matching Task (NMT). There was also two versions of the number matching task (NMT) called the NMT squares and NMT 4s, where each stimulus was either a large rectangle made up of smaller numbers, or a large "4" made up of smaller numbers, respectively (Figure 2). Each version of the NMT required the participant to ignore all 0s and 5s (in the same way as ignoring blue and green in the CMT), while attending to the numbers of interest (1, 2, 3, 4, 6, 7, 8). NMT control blocks included stimuli comprising only 0s and 5s, which are not numbers of interest. All of the other rules and details from the CMT applied to the NMT, substituting numbers in lieu of colours.

Akin to the CMT balloon, the NMT squares was a facilitating task, because the small numbers that make up the rectangle are arranged symmetrically, which makes visual identification of each small number relatively easy. The M-demand for this task was modelled as n + 1, where n is the number of numbers of interest, and 1 refers to the level of executive demand. The NMT 4s was a misleading task, because the spatial organization of the numbers was asymmetrical, which made visual identification of the small numbers more difficult. Further, the global shape of a 4 provided an additional source of interference (akin to the face of the clown in the CMT clown, which was to be ignored), requiring suppression from the I-operator in order to complete the task successfully. The M-demand of this task was modelled as n + 2, where the misleading nature of this task resulted in an executive demand level of 2, one unit higher than the NMT squares.

Optimization of the Current Study. The current study optimized the CMT and NMT for fMRI in order to maximize the quality and power of fMRI data. The current paradigm differed from the original paradigm (Arsalidou et al., 2013) in the following ways: (1) A considerable increase in fixation time between blocks was added in order to provide more time for task-related brain activation to return to baseline levels, allowing for a more accurate and reliable estimation of baseline BOLD signal across each run, thus yielding more accurate and reliable estimates of task-related BOLD signal responses. (2) The control stimuli were simplified so that there was only one class of control stimuli, which was matched the number of stimulus presentations for each difficulty level. (3) The order of block presentation was changed to be randomized for each participant. (4) fMRI scanning parameters were optimized to yield the highest spatial resolution possible with whole-brain coverage at the highest possible temporal resolution.

Practice Tasks. There were short practice versions of the CMT balloon, CMT clown, and NMT 4s that were used to train participants, consisting of 4 task blocks of 4 stimuli each, where one block was a control block. The structure and timing of the practice tasks were intended to prepare the participant for the real tasks. There was no shortened version of the NMT squares because the full version was completed by the participant during the practice session. The NMT squares was not included in the MRI session because it would have made the total time of the MRI session, a total of two hours, too long to be completed at once by participants. The stimuli used in each practice task were orthogonal to the stimuli used in the real MRI tasks.

Procedure

Practice Session. Initially, new participants were trained in the MRI simulator. Task instructions were administered with PowerPoint (www.microsoft.com) while the participant was approximately 60 cm away from the screen. The participant completed a motion reduction training session, consisting of a series of instructed movements in the simulator, while observing their head motion on a screen, followed by an instructional script of how to stay still in the real

Next, the participant completed 3 practice tasks, which were shortened versions of the CMT balloon, CMT clown, and NMT 4s followed by the full length NMT squares. The participant made responses on the MRI simulator button box; there were simulator sounds playing, which are audio recordings of actual MRI sequences; and a replica of the 32-channel head coil to be used in the MRI scanner was placed over the participant's head. Participants viewed a projection screen via a mirror mounted to the head coil, as per the usual MRI setup. Accuracy feedback was provided to the participant after each practice task, and there was an opportunity for the participant to clarify task instructions.

MRI Session. The participant returned on a subsequent day (approximately one week later, depending on scheduling factors) to undergo the MRI session. First, task instructions were reiterated to the participant with a script, and the participant was screened for compatibility with the MRI environment. Once in the MRI scanner, the participant began with the CMT balloons, then completed the CMT clowns. An anatomical scan was administered next, followed by the NMT 4s. Participants were debriefed and paid $20.00 as compensation following the MRI scan. Finally, participants later completed a post-task survey online about the strategies they used to complete the tasks.

MRI Data Acquisition

Brain imaging data were collected with a 3T Siemens Magnetom Tim Trio scanner using a 32-channel head coil at the York MRI Facility. Anatomical images were acquired with a 3D magnetization prepared rapid gradient echo sequence (TR = 2300ms; TE = 2.62ms; 9° Flip angle; 1.0 mm isotropic voxels). A total of 12 task-runs were collected using BOLD functional scanning (TR = 2500 ms; 132 measurements per run; voxel size = 3.0 mm isotropic; 41 interleaved slices covering the whole brain, including the cerebellum).

Behavioural responses during the task-runs were collected using a 4-key MRI compatible button box designed for the right hand. Shim cushions were placed between the participant's head and the head coil to minimize head movement. Physiological recordings of heart rate and respiration were recorded using a pulse and respiration monitor that is interfaced with a biopac receiver box (www.biopac.com). Presentation (Neurobehavioural Systems Inc.) was used to control the presentation of stimuli and record behavioural responses.

Behavioural Analysis

Accuracy. Accuracy was calculated separately for each difficulty level within each task. In the current study, accuracy was calculated as the proportion of correct responses across each difficulty level that involved both same and different trials together. A participant's M-capacity was determined by the highest level of difficulty that they passed with an accuracy cut-off of 70% correct, plus the number of executive demand units according to the task type, which has been the protocol used in previous research using this paradigm (Arsalidou et al., 2010; Arsalidou et al., 2013). For example, If a participant completed up to difficulty level five in the CMT clown task, then their M-capacity would be 5 + 2, where 5 refers to the difficulty level or number of elements that needed to be attended to at once and 2 refers to the level of executive demand in the CMT clown task as defined by the TCO. The NMT 4s task had an executive demand of 2, equal to the CMT clown while the CMT balloon had an executive demand of 1. If a participant failed only one difficulty level, but passed another more difficult level, their M-capacity was the highest difficulty level passed, despite failing one lower difficulty level, which did occur in this dataset. A repeated-measures analysis of variance (ANOVA) was conducted with two factors, difficulty level and task type, in order to determine whether different tasks yield different measures of M-capacity, and whether or not there is an interaction between difficulty level and task type. Then, 3 separate one-way ANOVAs were calculated, one for each task, to assess the main effect of difficulty level within each task separately. Pairwise comparisons were calculated using the Bonferroni correction for multiple comparisons, as well as Tukey's HSD mean differences, in order to determine which differences between difficulty levels were driving the main effect of difficulty within each task. Mean accuracy and reaction time in each task were plotted against difficulty level, indexed by the number of elements to be attended to.

Reaction Time. Reaction time was calculated as the amount of time between each stimulus onset and the participant's first button response to that stimulus. A repeated measures ANOVA was calculated on the reaction time data to see if there was an interaction between task type and difficulty level. Three separate ANOVAs were also calculated to assess the effect of difficulty level within each task and pairwise comparisons were calculated to break down specific main effects of difficulty level on reaction time.

MRI Data Analysis

Preprocessing. All imaging data were preprocessed and analyzed using AFNI (Cox, 1996). The first 3 volumes of every run were automatically discarded during the image acquisition stage to account for signal intensity equilibration. Preprocessing of fMRI data for each participant began with despiking the BOLD signal time series, which replaces abrupt and extreme amplitude spikes with an interpolated estimate, temporally smoothing the BOLD signal time series at each voxel. Next, the data were corrected for motion, by aligning subsequent functional images within a run to the initial base image of each run. Then, spatial shifts between slices were corrected using slice-timing correction, and lastly, each functional data set was aligned with its corresponding anatomical image. The data were spatially normalized to Montreal Neurological Institute (MNI) space and spatially smoothed using a 6 mm full width at half maximum Gaussian filter. Finally, the data were scaled so that the time series of each voxel would have a mean of 100.

fMRI Individual Participant Analysis. The data from each participant within each task were first analyzed on an individual level using the 3dDeconvolve program in AFNI. BOLD signal associated with accurate blocks ($\geq$ 70% correct) of each difficulty level were regressed as variables of interest, while failed blocks of each difficulty level were regressed as separate nuisance variables (e.g. D1, D2, errorD1, errorD2 etc.). The BOLD signal associated with control blocks was also accounted for and regressed as its own variable. Statistical maps were generated for each participant, which quantified task-related BOLD signal associated with each difficulty level (made up of accurate blocks only) by subtracting from it the signal associated with control blocks, to yield 6 statistical maps for each participant for each task representing each level of difficulty for each task (e.g. balloonD1 – control, balloonD2 – control, balloonD3 – control, etc.).

fMRI Group Analysis

Individual participant statistical maps of each difficulty level (made up of accurate blocks), minus the control blocks, were introduced into a mixed effects group general linear model using the 3dNAOVA3 type 4 program in AFNI. The factors of task type (CMT balloon, CMT clown, and NMT 4s) and difficulty level (1-6) were treated as fixed effects, because their effects were expected to be the direct result of the manipulation of these independent factors, while the third factor of participant was treated as a random effect, because an effect related to particular participants is expected to be generalized to the greater population of healthy young adults. The interaction effect between task type and difficulty level was calculated in order to identify voxels that were differentially modulated across difficulty levels depending on task type. A linear trend contrast across difficulty levels was calculated within each individual task type (CMT balloons, CMT clowns, NMT 4s), and collapsed across tasks, in order to identify voxels which significantly fit the pattern: D1 < D2 < D3 < D4 < D5 < D6 (i.e., showed a monotonic increase or decrease in activity with increasing difficulty). Previous research (Van Snellenberg et al., 2015) suggesting a curvilinear relationship between cognitive load and BOLD signal motivated the calculation of a quadratic contrast, collapsed across tasks, in order to identify voxels that fit the pattern: D1 < D2 < D3 = D4 > D5 > D6 (i.e., showed a pattern of an initial increase, followed by a decrease, in activity with increasing difficulty, or vice versa).

ROI Analysis: Coordinates from Arsalidou and Colleagues (2013). In order to replicate the methods of Arsalidou and colleagues (2013) in an a priori analysis, for each participant, the average BOLD signal was extracted from each region of interest (ROI) corresponding with the coordinates from Arsalidou et al. (2013). The researchers originally derived these coordinates (Table 13) from peak activations in a linear contrast across each difficulty level of the CMT clown. First, a sphere with a 6 mm diameter was created around each set of ROI coordinates, then, all ROI spheres were combined into one map where each sphere was assigned its own numeric value. Next, the average BOLD signal associated with accurate blocks from each difficulty level was extracted using the 3dROIstats program in AFNI. Time series data for each participant across each difficulty level were averaged to yield the mean BOLD signal value for each difficulty level in each ROI separately. This procedure was repeated for all 3 tasks and the group mean BOLD signal values for each task were plotted against difficulty level.

ROI Analysis: Coordinates from Significant Clusters. In order to replicate the results of Arsalidou and colleagues (2013) in an a posteriori analysis, the locations of peak activation of significant clusters in the linear contrast (Figure 7), collapsed across tasks, conducted on the current fMRI data were identified (Table 15). Then, the same procedures described above were used to create ROIs and an ROI map. These ROIs were then used to extract the mean BOLD signal from each participant across all 6 levels of difficulty (made up of accurate blocks only) in each task. Finally, the group mean BOLD signal values were plotted against difficulty level according to task type.

ROI Analysis: Coordinates from Yeo and Colleagues (2011). In order to gain an initial understanding of how key neural networks in the brain are behaving across the 3 tasks, we used validated ROI coordinates from the literature to extract the mean BOLD signal from each participant across all 3 tasks. The coordinates were taken from Yeo and colleagues (2011), which have been classified as key nodes in the default network, dorsal attention network, and frontoparietal control network. The results from these coordinates provided a completely unbiased account of how the activity of each network was modulated by the task conditions. The same procedure used to extract the average BOLD signal across participants outlined in the first analysis was used here for all 3 tasks and the group mean BOLD signal values were plotted against difficulty level according to task type.

Linear and Quadratic Contrast from Difficulty One to Five. The results of the first linear contrast and the ROI analysis, lead to a follow up linear trend contrast was calculated to identify voxels that linearly increased from difficulty levels one to five rather than one to six in an attempt to determine if the data were fitted differently than when the linear trend included the 6th level of difficulty. This contrast identified voxels which fit the pattern: D1 < D2< D3 < D4 < D5 (i.e., showed a monotonic increase or decrease in activity with increasing difficulty). In addition to the linear contrast, another quadratic contrast was also calculated across difficulty levels 1-5 in order to identify voxels which fit the pattern: D1 < D2 = D3 = D4 > D5 (i.e., showed a pattern of an initial increase, followed by a decrease, in activity with increasing difficulty, or vice versa).

Results

Behavioural Results

Accuracy. Figure 4a depicts the decrease in average accuracy as the number of elements to be compared increased. Figure 5a depicts the number of participants who accurately passed each difficulty level with at least 70% accuracy, indexed across tasks as the number of elements to be remembered. Figure 5b depicts the M-capacity, as defined by the TCO, of participants in the study. A two-way ANOVA was used to assess the effect of task type and difficulty on accuracy. Main effects of task type $F(2,522) = 8.62$, $p<0.001$, and difficulty $F(5,522)=11.77$, $p< 0.001$ were significant but the interaction between the two factors was not (Table 1). A separate one-way ANOVA was conducted within each task to assess the effect of difficulty on accuracy. There was a significant main effect of difficulty for the CMT balloon task $F(5, 174) = 2.87$, $p < .05$; the CMT clown task $F(5,174) = 5.59$, $p < 0.001$; and the NTM 4s task $F(5,174) = 4.12$, $p < 0.01$ (Table 2). The mean and standard deviation of accuracy in each task across each difficulty level are displayed in (Table 3).

Independent pairwise t-tests between difficulty levels for each task (Table 4) revealed that for the CMT balloon task, accuracy for difficulty level six was lower than difficulty level one, $p < .05$. For the CMT Clown task, accuracy for difficulty level six was lower than difficulty level one, $p < .001$ and two, $p < .001$ and difficulty five was lower than difficulty level one, $p< .01$ and two, $p < .05$. For the NMT 4s task, accuracy for difficulty level six was lower than difficulty level one, $p<.001$ and two, $p<.005$. The mean differences between the mean accuracy in each task was assessed using Tukey's HSD (Table 5). The mean accuracy for the CMT balloons was significantly higher than both the CMT clown and CMT 4s task, $p<.01$, while the CMT clown and NMT 4s task do not differ from each other.

Reaction Time. Figure 4b depicts the increase in reaction time as the number of elements increased. A two-way ANOVA was used to assess the effect of task type and difficulty on reaction time. Significant main effects of task type $F(2,14594) = 505$, $p<0.0001$, and difficulty $F(5,14594)=621$, $p< 0.0001$ were found (Table 7). A separate one-way ANOVA was conducted within each task to assess the effect of difficulty on reaction time. A significant main effect of difficulty was found for the CMT balloon task $F(5,14594) = 182$, $p < .0001$; the CMT clown task

$F(5,14594) = 220$, $p < 0.0001$; and the NMT 4s task $F(5,14594) = 222$, $p < 0.0001$ (Table 8). The mean and standard deviation of accuracy in each task across each difficulty level are displayed in Table 9.

Independent pairwise t-tests between difficulty levels for each task (Table 10) revealed that for all tasks, each difficulty level had a longer reaction time than the previous one such that: D1 < D2 < D3 <D4 < D5 < D6, $p < .0001$; with the exception of D1 < D2 for the CMT balloon task, $p < .05$ (Table 10). Mean differences in reaction time across task type revealed that each task had higher reaction time than the previous one such that: CMT balloon < CMT clown < CMT 4s, $p< .0001$.

fMRI Results

fMRI Individual Participant Analysis. Before the group general linear model (GLM) analysis, data from individual participants were checked for accuracy, motion artifacts, and structural neurological abnormalities that may affect group statistical results. There was one participant who displayed abnormal structural characteristics throughout the white matter. These abnormalities were wide spread and believed to possibly affect the interconnections among brain regions; this participant was excluded from any further group analysis as a result.

There was a total of six participants who lacked a minimum of one task block per each condition that was completed with over 70% accuracy. Given the nature of our analysis, in that we intended to analyze accurate blocks only (in keeping with the 2013 procedure of Arsalidou and Colleagues), we elected to exclude these participants from any further group analysis. In accordance with these exclusions, the number of participants included in each group analysis was n= 23.

fMRI Group Analysis

Figure 18 shows the number of blocks completed with over 70% accuracy that were included in the following group analyses. A 3dANOVA was conducted with 23 participants, in order to assess whether there was an interaction between task type and difficulty level. The minimum cluster size was set to 20 voxels, $q < .001$ (FDR corrected). The optimized task design in the current investigation increased the amount of statistical power, allowing for a more conservative error threshold to be upheld for all of the contrasts of interest. Figure 6 displays three significant clusters: bilateral fusiform gyri, and the right middle occipital cortex, that showed an interaction between task type and difficulty level. The coordinates of peak activation for each significant cluster showing this interaction was identified and the mean BOLD signal time course was extracted from a 6mm diameter spherical ROI around the peak in each of these clusters. The mean BOLD time course in each ROI across participants was plotted against difficulty level according to task type (Figure 6). In order to identify brain regions that were modulated monotonically (in keeping with the Arsalidou and colleagues (2013) investigation) as a function of difficulty level in each task, a linear contrast of difficulty levels one to six was performed on the group fMRI data incorporating all three tasks. The linear contrast revealed an opposite pattern of activation and deactivation associated with increasing difficulty than expected and observed in previous work (Arsalidou et al., 2013). Namely, key default network areas, such as the medial PFC, anterior cingulate, and posterior cingulate, showed a positive linear effect (Figure 7, warm colours), implying an increase in activation with increasing difficulty. Conversely, key dorsal attention network (e.g., superior parietal lobes) and frontoparietal control network (e.g., bilateral middle frontal gyri) regions showed a negative linear effect, (Figure 7, cool colours), implying a decrease in activation with increasing difficulty (Figure 7).

Quadratic Contrast of Difficulty Level One to Six. The investigation of Arsalidou and colleagues (2013) using the CMT clown task suggested that a linear relationship between BOLD signal and difficulty level existed, such that an increase in difficulty resulted in a linear increase in attention and control areas, and conversely , a linear decrease (i.e., increased suppression) in default network areas. There was, however, a visible trend in the previous results of Arsalidou et al. (2013) showing that BOLD signal either plateaued between difficulty levels five and six, or in some brain regions, showed a reversal of the linear trend direction – i.e., decreased slightly in attention/control network areas and increased in default network areas – suggesting a curvilinear (i.e., quadratic) trend, rather than a purely linear trend. Considering that the previous study included only 10 participants, it is possible that this visible trend was not significant due to a lack of statistical power. Given that there were more than double the number of participants and triple the number of task runs collected in the current study, detecting a significant curvilinear trend might be expected.

This motivated an additional investigation which was a quadratic contrast of difficulty level one to six to identify voxels that fit the trend: D1<D2<D3=D4>D5>D6 (Figure 8). This quadratic contrast flipped the pattern of activation found in the Linear contrast (Figure 7) by displaying key default network regions in cool colours, signifying that they negatively related to the quadratic trend, and displaying key frontoparietal and dorsal attention network regions in warm colours indicating that they positively related to the quadratic trend (Figure 8).

ROI Analysis: Coordinates from Arsalidou and Colleagues (2013). Given the observation of an unexpected pattern of activation (Figures 7 and 8) that was not in line with the previous investigation by Arsalidou and colleagues (2013), a closer look at the fluctuation of BOLD signal was required to gain an understanding of how the brain was being modulated by difficulty level. In order to determine how BOLD signal was modulated in different brain regions as a function of difficulty, and to replicate previous findings, ROI coordinates were taken from Arsalidou and colleagues (2013). These coordinates (Table 13) were based on the peak activations in a linear contrast of difficulty level calculated on fMRI data collected during the CMT clown task (Arsalidou et al., 2013). In the current study, 6mm diameter spherical ROIs centred on these coordinates were used to extract the mean BOLD signal from each participant included in the linear contrast analysis (n= 23). The average BOLD signal at the individual participant level was extracted for each difficulty level. Figures 10 and 11 show the mean BOLD signal across participants plotted against difficulty level for select ROIs affiliated with the: default and frontoparietal control networks. Voxels that were significantly modulated by difficulty in the linear contrast (Figure 7) are shown in the activation maps (Figures 10 and 11), and the crosshairs indicate where the ROI coordinates overlap with the statistical map from the linear contrast (Figure 7). The ROI graphs show BOLD signal change increasing with difficulty, then a sharp change of direction that occurs between four and five elements. This change of direction was replicated in all of the other sets of ROIs and often happens earlier for the clown task than the balloon task, which according to the TCO, is due to the difference in executive demand across a facilitating and misleading task. This directional change of BOLD signal was presumed to be the source of misfit for the original linear contrast of difficulty levels one to six (Figure 7). A one-way ANOVA was conducted to assess the main effect of difficulty on BOLD signal in each ROI; Table 8 displays corresponding test-statistics for each ROI in which BOLD signal was significantly modulated by difficulty level, $p < .05$.

ROI Analysis: Coordinates from Significant Clusters. In order to replicate the procedure of Arsalidou and colleagues (2013) in a data driven way within the current study, the mean BOLD signal was extracted from ROIs created around peak activation sites present in the first linear contrast (Figure 7d). Mean BOLD signal was extracted from select ROIs associated with each of the three networks and plotted against difficulty level for each task, in order to depict the modulation of BOLD signal by difficulty level (Figures 12, 13, and 14), and a one-way ANOVA was conducted for each ROI (Table 11). The BOLD signal within each ROI was significantly modulated by difficulty level $p < .0001$.

ROI Analysis: Coordinates from Yeo and Colleagues (2011). To investigate in a completely unbiased way the BOLD signal time courses in well-established critical nodes of the default, frontoparietal, and dorsal attention networks, ROIs were created using coordinates (Table 12) from Yeo et al. (2011). The mean BOLD signal across participants was extracted from these ROIs and plotted against difficulty level for each task (Figures 15, 16, and 17). Multiple one-way ANOVAs were conducted to assess the effect of difficulty on BOLD signal in each ROI; corresponding test statistics are listed in Table 13. Every ROI was significantly modulated by difficulty level $p < .0001$, except for in the parahippocampal complex and the dorsal anterior PFC which did not overlap with significant voxels in the linear contrast statistical map (Figure 17).

Linear and Quadratic Contrast from Difficulty One to Five. The replicated pattern across the ROI graphs suggested that the relationship between difficulty and BOLD signal may be a monotonic linear relationship from difficulty 1-5. In order to test this, another linear contrast across only difficulty levels 1-5 (i.e., excluding level 6) was conducted to determine if a reversal of the linear trend of a activation at the highest difficulty level may have been the source of misfit for the original linear model (Figure 7).

The linear contrast of D1 < D2 < D3 < D4 < D5 produced a statistical map (Figure 9a) that was the inverse of that found for the linear contrast that included all 6 difficulty levels (Figure 7), and closely replicated the pattern reported by Arsalidou et al. (2013), and thus, was consistent with traditional expectations of default, dorsal attention, and frontoparietal control network behaviour. Default network areas (e.g., PFC and posterior cingulate) decreased across difficulty (Figure 9a, cool colours). Key dorsal attention network areas (e.g., bilateral superior parietal lobes) and frontoparietal control network areas (e.g., bilateral middle frontal gyri) increased in activation from difficulty level one to five (Figure 9a, warm colours). For completeness, we also ran a quadratic contrast on difficulty level one to five (Figure 9b) in order to determine that the clusters found in the linear contrast were indeed behaving linearly and not simply showing a pattern being driven by one condition. The quadratic contrast (Figure 9b) yielded no significant voxels, indicating that the clusters we found in the linear contrast of difficulty one to five (Figure 9a) were indeed linearly modulated.

Discussion

The objective of the present study was to begin an investigation of the neural underpinnings of M-capacity within the TCO framework, and in the context of two WM tasks that have not yet been studied with neuroimaging. It was hypothesized that 1) the patterns of activation and deactivation observed by Arsalidou and colleagues (2013) would be replicated and 2) that there would be a significant interaction effect between task type and set size, marking an effect of the (I)-operator in the brain. Our observations have ultimately informed the existing TCO and motivated a number of future directions that would further investigate the interplay between key neural networks and M-capacity in healthy young adults; with the benefit of using a paradigm that can be used to test children and older adults for future comparisons of the trajectory of M-capacity across the lifespan.

Summary of Results

Behavioural Results. The relationship between accuracy and difficulty level observed was an inverse linear relationship, where accuracy decreased as difficulty increased, replicating previous observations (Arsalidou et al., 2013). Furthermore, reaction time increased as difficulty increased, in line with Arsalidou and colleagues (2013). Behaviourally, there was no significant interaction between task type and difficulty level for task accuracy, however, a main effect of task type leading to post hoc analysis of mean differences revealed that accuracy did not differ between the CMT clown and NMT 4s, but was higher in the CMT balloon than both of the other tasks. This suggested that the additional recruitment of the I-operator in the CMT clown and NMT 4s resulted in a comparable increase in executive demand relative to the CMT balloon while varying M-demand (set size).

fMRI Results. The bilateral fusiform gyri and the right middle occipital gyrus displayed a significant interaction effect between task type and difficulty level. This suggested that these particular brain regions were modulated by difficulty level differently depending on which task was being completed. Notably, the average BOLD signal in the left fusiform gyrus showed a lower amplitude in the NMT 4s task compared to the CMT balloon and CMT clown. There is strong evidence implicating the left fusiform gyrus as a key contributor in the holistic perception of letters as words (McCandliss et al., 2003). Druzgal and D'Esposito (2001) studied fusiform activation during an n-back WM task using face stimuli, and found that it was directly modulated by cognitive load. The CMT clown includes a face in the stimuli but the CMT balloon and NMT 4s do not. Given that numbers are very different from words and faces, the left fusiform gyrus is an important region for conducting further analysis to determine what it is about the NMT 4s task that is driving this interaction because the current investigation was not able to answer this question per se.

Arsalidou and colleagues (2013) found an increase in BOLD signal within brain regions associated with the frontoparietal control network. They also found a decrease in BOLD signal within brain regions associated with the default network. The current investigation intended to replicate and extend these findings by preforming a linear contrast on fMRI data collected during the completion of the CMT balloon, CMT clown, and NMT 4s. The first linear contrast was intended to identify voxels in which BOLD signal increased as a function of difficulty levels one to six (difficulty levels are indexed by number of elements for the sake of comparing across tasks). A pattern of activation that was the complete inverse of that found by Arsalidou and colleagues (2013) was found for the first linear contrast (Figure 7), which indicated that the BOLD signal did not relate to difficulty in a linear fashion. A visible but not significant curvilinear trend in BOLD signal across difficulty level in the investigation by Arsalidou and colleagues (2013) motivated a quadratic contrast to be conducted to identify voxels which displayed a U-shape function (inverted or upright) in relation to the factor of difficulty. This contrast (Figure 8) yielded a statistical map that was the inverse of the pattern produced by the original linear contrast analysis (Figure 7), entirely consistent with the pattern of activation found by Arsalidou and Colleagues (2013), and also in line with traditional expectations of how the default and frontoparietal control networks typically behave in externally driven cognitive tasks (Raichle et al., 2001; Buckner et al., 2008; Fox et al., 2005; Corbetta et al., 2008). Taken together, these observations suggested that the relationship between M-demand and M-capacity, while monotonically linear in behavioural data, may not be monotonically linear in the brain, which is in line with a more recent investigation by Van Snellenberg and colleagues (2015). Notably, Van Snellenberg and colleagues (2015) aimed to investigate how BOLD signal was modulated by cognitive load beyond an individual's WM capacity limit, even though behavioural performance was still accurate. Therefore, the linear (Figure 7) and quadratic (Figure 8) contrast results may have been driven by a change in how the brain handles higher levels of M-demand.

In order to gain a closer understanding of how the BOLD signal was modulated by cognitive load, the average BOLD signal was extracted in three different sets of coordinates, each with the purpose of investigating neural network behaviour from a different experimental standpoint. The first set of coordinates was taken from Arsalidou and colleagues (2013) with the intention of replicating the BOLD signal modulation previously observed. The ROI graphs showed an increase/decrease in BOLD signal for the frontoparietal control network and default network, respectively, with a sharp reversal in the direction of activation around difficulty level four or five often depending on the particular task. The CMT balloon often yielded a directional change in BOLD amplitude at a set size one step higher than the other two tasks, which was attributable to the additional engagement of the I-operator in the misleading tasks. This sharp directional change was foreshadowed by the change in the pattern of activation between the linear and quadratic contrast of difficulty one to six.

The second set of coordinates was taken form peak activations in the first linear contrast of difficulty one to six with the intention of replicating the procedure of Arsalidou and colleagues (2013) in a data driven way using the current results. These ROI graphs were consistent with those of the previous ROI coordinates, with the addition of key dorsal attention network regions. A third set of a priori coordinates was taken from the literature (Yeo et al., 2011). These literature coordinates have been validated as key functional network nodes of the default, dorsal attention, and frontoparietal control networks. Again, the corresponding ROI graphs display a similar pattern to the previous two sets of coordinates.

Upon investigation of the modulation of BOLD signal in key ROIs across difficulty level, additional linear and quadratic contrasts were conducted across difficulty levels one to five only (i.e., excluding difficulty level 6) to identify voxels that showed a monotonic increase/decrease in BOLD signal at difficulty levels less likely to have exceeded the participant's M-capacity. Given the ROI graphs, it was possible that difficulty level six was the source of misfit in the previous linear contrast (Figure 7) and if this was the case, then omitting it should produce a statistical map consistent with that reported by Arsalidou and colleagues (2013). The reversal of the relationship between BOLD signal and difficulty level at higher levels of M-demand in the present investigation suggested that the brain was doing something that is computationally different at a high cognitive load relative to low cognitive load in order to successfully complete the task at hand. In line with this rationale, the linear contrast across difficulty levels one to five yielded a statistical map that was the inverse of that found when difficulty level 6 was included in the analysis, and was therefore consistent with expectations of how the default, dorsal attention, and frontoparietal control networks normally behave (Raichle et al., 2001; Buckner et al., 2008; Fox et al., 2005; Corbetta et al., 2008). Furthermore, the quadratic contrast across only difficulty levels one to five yielded no significant activation, meaning that there were no voxels associated with difficulty one to five that displayed a U-function in relation to the factor of difficulty level.

Conclusions

The first hypothesis that the patterns of activation and deactivation observed by Arsalidou and colleagues (2013) would be replicated was confirmed but with additional insight. While default, dorsal attention, and frontoparietal control network regions did show the initial linear pattern observed by Arsalidou and colleagues (2013), the visible trend of a reversal of this linear relationship was observed as a significant effect in the present investigation. The second hypothesis that there would be an interaction between task type and difficulty level was also confirmed in that there were three clusters in the brain significantly showing this interaction.

Taken together, these observations suggested that M-capacity in the brain related to difficulty in a monotonic way between difficulty levels one to five, but once set size approached six elements, there was a reversal of the patterns of activation across the whole brain where default network areas increased in activation, and dorsal attention and frontoparietal network areas decreased in activation. These observations suggested a dynamic interplay of these three neural networks where the brain must approach a task with high M-demand differently than a task with low M-demand in order to maintain accurate performance. The particular location of this amplitude reversal (i.e. at difficulty four or five), which often happened at a higher cognitive load in the CMT balloon relative to the CMT clown and NMT 4s can be attributed to the difference in additional engagement of the I-operator where an increase in executive demand results in an earlier neurological switch from a low to a high cognitive load computation strategy.

Multiple contrasts assessing the linear relationship between BOLD response and difficulty suggested that a simple linear contrast does not fit the data well and that more sophisticated analysis should be conducted in order to clearly understand how large-scale functional networks in the brain underlie M-capacity, as the present observations suggest a dynamic and complex interplay between M-demand, executive demand, and the brain areas that serve WM function. These results will lead to a more complex multivariate analysis of the neural underpinnings of M-capacity (see Future Directions).

Limitations

With respect to the current study, the ability to gain a clear understanding of how M-capacity is carried out in the brain is limited. This study used a univariate approach to initially investigate how the brain is modulated by cognitive load in a WM task. While it was determined that the NMT 4s had a comparable executive demand to the CMT clown, the interaction effect suggests that within the left fusiform gyrus, there is a difference in the degree to which BOLD signal is modulated by the NMT 4s specifically. The current analysis was not able to directly answer the question of why the NMT 4s differed from the CMT balloon and clown even though executive demand was controlled between the NMT 4s and CMT clown.

The current investigation also did not tell us how the default, dorsal attention, and frontoparietal control networks were connected within and across neural networks as cognitive load increased. Furthermore, the current investigation was not able to look into the nature of Individual differences and gender differences in M-capacity, and if these differences do exist, which factors may drive these differences.

Future Directions

In order to address these limitations, continued analysis of this data set is planned with a number of multivariate techniques that will help clarify how M-capacity is instantiated in the brain. Firstly, a task-based functional connectivity analysis will be conducted using seed PLS in order to optimally account for whole-brain patterns of activity that covary with a particular seed region corresponding to key nodes in each network of interest. The patterns of covariance can be quantified across levels of cognitive load and this will allow for insight into the degree to which the frontoparietal control network is functionally related to the default and dorsal attention networks.

In order to further investigate the effect of task type, the left fusiform gyrus activation can be used as a seed region, and seed PLS can be used to identify whole-brain patterns of covariance specifically between the left fusiform gyrus and the rest of the brain. This will help to determine whether or not the difference between the CMT and NMT in the left fusiform is driven by a difference in stimuli or not depending on how it connects to other sensory brain areas.

Behavioural PLS can be used to account for whole-brain patterns of covariance in BOLD signal correlated with task accuracy, which will determine which patterns of brain activation are associated with certain levels of behavioural performance, leading to an understanding of individual differences in the neural underpinnings of M-capacity.

References

Arsalidou, M., Pascual-Leone, J., & Johnson, J. (2010). Misleading cues improve developmental assessment of working memory capacity: the color matching tasks. Cognitive Development, 25(3), 262-277.

Arsalidou, M., Pascual-Leone, J., Johnson, J., Morris, D., & Taylor, M. J. (2013). A balancing act of the brain: Activations and deactivations driven by cognitive load. Brain and Behavior, 3(3), 273-285.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes1. In Psychology of learning and motivation (Vol. 2, pp. 89-195). Academic Press.

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and WM. Neuroscience, 139(1), 201-208.

Baddeley, A. (2003). WM: looking back and looking forward. Nature reviews neuroscience, 4(10), 829.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. Annals of the New York Academy of Sciences, 1124(1), 1-38.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., ... & Weinberger, D. R. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. Cerebral cortex, 9(1), 20-26.

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific WM systems. Memory & Cognition, 30(7), 1086-1095.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature reviews neuroscience, 3(3), 201.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. Neuron, 58(3), 306-324.

Crowder, R. G. (1982). The demise of short-term memory. Acta Psychologica, 50(3), 291-323.

D'Esposito, M., & Postle, B. R. (1999). The dependence of span and delayed-response performance on PFC. Neuropsychologia, 37(11), 1303-1315.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in WM and reading. Journal of verbal learning and verbal behavior, 19(4), 450-466.

Druzgal, T. J., & D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of WM load1. Cognitive Brain Research, 10(3), 355-364.

Engle, R. W. (2001). What is WM capacity? WM. Science, 255(5044), pp.556-559.

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive architecture of WM. Neuron, 88(1), 33-46.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences, 102(27), 9673-9678.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral PFC. Journal of neurophysiology, 61(2), 331-349.

Fuster, J. M. (2009). Cortex and memory: emergence of a new paradigm. Journal of cognitive neuroscience, 21(11), 2047-2072.

Gazzaley, A., Rissman, J., & D'esposito, M. (2004). Functional connectivity during WM maintenance. Cognitive, Affective, & Behavioral Neuroscience, 4(4), 580-599.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. Journal of experimental psychology, 55(4), 352.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. Trends in cognitive sciences, 7(7), 293-299.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial WM following frontal lobe lesions in man. Neuropsychologia, 28(10), 1021-1034.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back WM paradigm: A meta-analysis of normative functional neuroimaging studies. Human brain mapping, 25(1), 46-59.

Pascual-Leone, J., & Baillargeon, R. (1994). Developmental measurement of mental attention. International Journal of Behavioral Development, 17(1), 161-200.

Pascual-Leone, J., & Johnson, J. (2005). A dialectical constructivist view of developmental intelligence. In O. Wilhelm & R. Engle (Eds.), Handbook of understanding andmeasuring intelligence (pp. 177-201). Thousand Oaks, CA: Sage.

Pennings, A. H., & Hessels, M. G. (1996). The measurement of mental attentional capacity: A neo-Piagetian developmental study. Intelligence, 23(1), 59-78.

Raichle, M. E. (2010). Two views of brain function. Trends in cognitive sciences, 14(4), 180-190.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of neurology, neurosurgery, and psychiatry, 20(1), 11.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. Journal of cognitive neuroscience, 9(5), 648-663.

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. Journal of cognitive neuroscience, 21(3), 489-510.

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. Neuroimage, 53(1), 303-317.

Spreng, R. N., Stevens, W. D., Viviano, J. D., & Schacter, D. L. (2016). Attenuated anticorrelation between the default and dorsal attention networks with aging: evidence from task and rest. Neurobiology of aging, 45, 149-160.

Ungerleider, S. K. A. L. G. (2000). Mechanisms of visual attention in the human cortex. Annual review of neuroscience, 23(1), 315-341.

Van Snellenberg, J. X., Slifstein, M., Read, C., Weber, J., Thompson, J. L., Wager, T. D., ... & Smith, E. E. (2015). Dynamic shifts in brain network activation during supracapacity WM task performance. Human brain mapping, 36(4), 1245-1264.

Tables

Table 1: Two-Way ANOVA on the effect of task type and difficulty on Accuracy

	Df	Sum sq	mean sq	*F*	*p*	*Significance*
balloon	5	5271	1054.2	2.868	0.0163	*
clown	5	11027	2205.4	5.594	8.32E-05	***
4s	5	7719	1543.8	4.115	0.00148	**

Table2: Three One-Way ANOVAs effect of Difficulty on Accuracy in each task

	Df	Sum sq	Mean sq	*F*	*p*	*Significance*
task	2	6538	3269	8.625	0.00021	***
difficulty	5	22312	4462	11.774	8.36E-11	***
Interaction	10	1706	171	0.45	0.9211	
Residuals	522	197834	379			

Table 3: Mean and Standard Deviation of Accuracy across difficulty levels within each task

Number of Elements	Balloon		Clown		4s	
	Mean	SD	Mean	SD	Mean	SD
1	83.9	18.4	80.2	24	74.9	22.3
2	79.9	20.7	78.1	22.5	76.7	19.6
3	81.2	17.1	75	22	70.2	19.4
4	76.2	22.1	70.4	19.8	69.5	19.2
5	74	19.8	61.5	13.1	65.8	17.9
6	67.4	16.2	59.8	15.5	56.7	17.4

Table 4: P-values of pairwise Comparisons of accuracy across difficulty level Using t-test with Bonferroni Correction, *df* = 22.

All Tasks	D1	D2	D3	D4	D5
D2	1				
D3	1	1			
D4	0.13751	0.5248	1		
D5	0.00033	0.00261	0.06904	1	
D6	1.00E-08	1.80E-07	2.40E-05	0.00397	0.68121
Balloon					
D2	1				
D3	1	1			
D4	1	1	1		
D5	0.712	1	1	1	
D6	0.015	0.187	0.088	1	1
Clown					
D2	1				
D3	1	1			
D4	0.8336	1	1		
D5	0.0053	0.0224	0.142	1	
D6	0.0014	0.0068	0.0506	0.6038	1
4s					
D2	1				
D3	1	1			
D4	1	1	1		
D5	1	0.4747	1	1	
D6	0.0053	0.0014	0.1098	0.1647	1

Table 5: Mean Differences of accuracy between tasks using Tukey's HSD

	MD	*p*
balloon-4s	8.13492	0.0002471***
clown-4s	1.86508	0.6349287
clown-balloon	-6.2698	0.006669**

Table 6: Mean Difference of accuracy between levels of difficulty across each task

	Balloon		Clown		4s	
	MD	*p*	*MD*	*p*	*MD*	*p*
D2-D1	-4.0476	0.9639647	-2.1429	0.9983475	1.78571	0.9992276
D3-D1	-2.7381	0.9937688	-5.2381	0.9101557	-4.6429	0.9387071
D4-D1	-7.7381	0.6238459	-9.8809	0.3891056	-5.3572	0.8921673
D5-D1	-9.881	0.3488064	-18.69	0.0046757**	-9.0476	0.462549
D6-D1	-16.548	0.0128092*	-20.476	0.001326**	-18.214	0.0047309**
D3-D2	1.30952	0.9998219	-3.0952	0.9906562	-6.4286	0.7926372
D4-D2	-3.6905	0.9758564	-7.7381	0.6587599	-7.1429	0.7097619
D5-D2	-5.8333	0.8466301	-16.548	0.0183656*	-10.833	0.2589985
D6-D2	-12.5	0.1225069	-18.333	0.0059399**	-20	0.0012993**
D4-D3	-5	0.9140998	-4.6429	0.9446526	-0.7143	0.9999916
D5-D3	-7.1429	0.7007394	-13.452	0.0969836	-4.4048	0.9506689
D6-D3	-13.81	0.0638929	-15.238	0.0389457	-13.571	0.0776381
D5-D4	-2.1429	0.998045	-8.8095	0.521605	-3.6905	0.9769133
D6-D4	-8.8095	0.4816015	-10.595	0.3097422	-12.857	0.1100742
D6-D5	-6.6667	0.7584296	-1.7857	0.9993156	-9.1667	0.447385

Table 7: Two-way ANOVA on reaction time

	Df	Sum Sq	Mean Sq	*F* value	*p*	*Significance*
Task	2	274	137.08	505.19	< .0001	***
Difficulty	5	843	168.51	621.014	< .0001	***
Interaction	10	16	1.55	5.727	<.0001	***
Residuals	14594	3960	0.27			

Table 8: Three one-way ANOVAs on reaction time

	df	Sum sq	Mean sq	*F*	*p*	*Significance*
balloon	5	198.3	39.65	182	<.0001	***
clown	5	311.7	62.33	220	<.0001	***
4s	5	348.2	69.63	222.9	<.0001	**

Table 9: Mean and Standard deviation of reaction time across tasks

Number of Elements	Balloon		Clown		4s	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
1	0.975	0.389	1.11	0.437	1.16	0.47
2	1.04	0.399	1.27	0.459	1.31	0.505
3	1.2	0.437	1.49	0.534	1.54	0.548
4	1.34	0.467	1.65	0.552	1.7	0.562
5	1.44	0.514	1.75	0.582	1.83	0.664
6	1.54	0.569	1.8	0.605	1.89	0.588

Table 10: *p* values of pairwise comparisons of reaction time across difficulty levels using Bonferroni Correction

df = 22.

All Tasks	D1	D2	D3	D4	D5
D2	1.50E-15				
D3	< 2e-16	< 2e-16			
D4	< 2e-16	< 2e-16	< 2e-16		
D5	< 2e-16	< 2e-16	< 2e-16	2.20E-10	
D6	< 2e-16	< 2e-16	< 2e-16	< 2e-16	4.00E-05
Balloon					
D2	0.04208				
D3	< 2e-16	1.20E-10			
D4	< 2e-16	< 2e-16	2.00E-07		
D5	< 2e-16	< 2e-16	< 2e-16	0.00026	
D6	< 2e-16	< 2e-16	< 2e-16	< 2e-16	0.00017
Clown					
D2	2.80E-08				
D3	< 2e-16	3.60E-16			
D4	< 2e-16	< 2e-16	1.50E-08		
D5	< 2e-16	< 2e-16	< 2e-16	3.00E-03	
D6	< 2e-16	< 2e-16	< 2e-16	3.50E-07	9.05E-01
4s					
D2	1.80E-07				
D3	< 2e-16	3.20E-15			
D4	< 2e-16	< 2e-16	1.30E-07		
D5	< 2e-16	< 2e-16	< 2e-16	3.50E-05	
D6	< 2e-16	< 2e-16	< 2e-16	1.10E-10	5.40E-01

Table 11: Mean differences of reaction time across task type using Tukey's HSD

Task	MD	*p*
clown-balloon	0.25982	0
fours-balloon	0.3144	0
fours-clown	0.05458	0.0000007***

Table 12: Mean differences of reaction time across difficulty levels using Tukey's HSD

	Balloon		Clown		4s	
	MD	*p*	*MD*	*p*	*MD*	*p*
D2-D1	0.0693	0.0334094	0.1599	0	0.15738	0.0000002
D3-D1	0.22809	0	0.38513	0	0.38474	0
D4-D1	0.36103	0	0.54609	0	0.54389	0
D5-D1	0.46152	0	0.64399	0	0.67564	0
D6-D1	0.56329	0	0.6935	0	0.73403	0
D3-D2	0.15879	0	0.22523	0	0.22735	0
D4-D2	0.29173	0	0.38619	0	0.38651	0
D5-D2	0.39222	0	0.48409	0	0.51825	0
D6-D2	0.49399	0	0.53361	0	0.57665	0
D4-D3	0.13294	0.0000002	0.16096	0	0.15916	0.0000001
D5-D3	0.23343	0	0.25886	0	0.2909	0
D6-D3	0.3352	0	0.30838	0	0.34929	0
D5-D4	0.10049	0.0002494	0.0979	0.0027183	0.13175	0.0000344
D6-D4	0.20226	0	0.14742	0.0000004	0.19014	0
D6-D5	0.10177	0.0001688	0.04952	0.4155237	0.05839	0.2897792

Table 13*:* ROI coordinates from Arsalidou and colleagues (2013)
Used to compare current fMRI data with previous study of the CMT clown task

Hemisphere	Peak Region	Brodman Area	MNI Coordinates (mm)		
			x	y	z
L	Medial Frontal Gyrus	10/32	-1	49	1
R	Medial Frontal Gurus	10	4	50	1
L	Anterior Cingulate	32	-3	43	1
R	Superior Temporal Gyrus	42	59	-26	17
R	Superior Temporal Gyrus	38	41	11	-20
L	Posterior Cingulate	31	-12	-52	24
R	PostCentral Gyrus	40	53	-23	17
R	Cingulate Gyrus	32	14	26	26
L	Cingulate Gyrus	32	-5	20	36
R	Inferior Frontal Gyrus	9	40	6	23
L	Inferior Frontal Gyrus	9	-39	8	28
R	Middle Frontal Gyrus	9	39	30	31
R	Middle Frontal Gyrus	10	34	45	24
R	Middle Frontal Gyrus	46	42	28	22
L	Middle Frontal Gyrus	46	-40	39	18
L	Middle Frontal Gyrus	10	-39	49	13
R	Middle Frontal Gyrus	6	31	4	47
R	Insula	13	32	20	1
L	Insula	13	-29	23	9
L	Fusiform Gyrus	37	-36	-56	-13
R	Fusiform Gyrus	39	25	-58	-9
R	Inferior Parietal Lobe	40	35	-48	41
L	Precentral Gyrus	6	-40	4	34
R	Precuneus	7	21	-61	43
L	Precuneus	7	-22	-60	42
R	Thalamus		7	-21	6
L	Declive1		-3	-63	-11
L	Declive2		-8	-84	-20
L	Declive3		-3	-86	-16
R	Cuneus	19	30	-83	33

Table 14: Effect of Difficulty on each ROI from Arsalidou et al., (2013)

ROI	*F* (Difficulty)	*p*	*Significance*
Medial Frontal Gyrus	51.917395	2.20E-16	***
Medial Frontal Gurus	49.4361435	2.20E-16	***
Anterior Cingulate	55.119107	2.20E-16	***
Superior Temporal Gyrus	28.5602268	0.01198	*
Superior Temporal Gyrus	2.9623193	<2e-16	***
Posterior Cingulate	41.0948256	2.00E-16	***
Postcentral Gyrus	28.734776	<2e-16	***
Cingulate Gyrus	20.5221815	0.04567	*
Cingulate Gyrus	2.2789589	<2e-16	***
Inferior Frontal Gyrus	30.0265743	2.00E-16	***
Inferior Frontal Gyrus	67.0978561	<2e-16	***
Middle Frontal Gyrus	47.0133721	<2e-16	***
Middle Frontal Gyrus	20.9709142	2.00E-16	***
Middle Frontal Gyrus	30.0808605	0.006392	**
Middle Frontal Gyrus	3.2737735	2.50E-08	***
Middle Frontal Gyrus	9.11906666	<2e-16	***
Middle Frontal Gyrus	51.4599188	2.00E-16	***
Insula	93.5153753	<2e-16	***
Insula	70.7952262	5.90E-12	***
Fusiform Gyrus	13.0187848	0.0007978	***
Fusiform Gyrus	4.280419	2.20E-16	***
Inferior Parietal Lobe	55.203065	0.002961	**
Precentral Gyrus	3.6498035	2.20E-16	***
Precuneus	24.433983	0.2421	
Precuneus	1.3492523	<2e-16	***
Thalamus	41.2374636	2.00E-16	***
Declive1	45.325882	2.00E-16	***
Declive2	66.0860766	2.20E-16	***
Declive3	43.1895073	<2e-16	***
Cuneus	18.5458135		

Table 15: Coordinates of peek activation from significant clusters
Derived from linear contrast across difficulty level. (Figure 7)

		MNI Coordinates (mm)			Cluster Size	
Peak Region	Hemisphere	x	y	z	(in voxels)	Network Affiliation
Angular Gyrus	R	54	-69	48	68	Default
Angular Gyrus	L	-42	-78	45	239	Default
Medial Frontal Gyrus	L	0	63	-3	985	Default
Superior Frontal Gyrus	R	33	0	66	88	Default
Superior Frontal Gyrus	L	-21	45	57	32	Default
Posterior Cingulate	L	-9	-54	12	20	Default
Superior Temporal Gyrus	L	-54	-30	21	62	Default
Superior Temporal Gyrus	L	-42	-12	0	28	Default
Superior Parietal Lobe	R	21	-78	60	1316	Dorsal Attention
Superior Parietal Lobe	L	-12	-81	57	1346	Dorsal Attention
Middle Frontal Gyrus	R	51	39	33	264	Frontoparietal
Middle Frontal Gyrus	L	-51	33	36	96	Frontoparietal
Insula	R	36	24	0	130	Frontoparietal
Insula	L	-30	27	3	102	Frontoparietal
Thalamus	R	9	-15	12	243	
Thalamus	L	-9	-15	12	159	
Middle Cigulate Cortex	L	0	-6	51	31	
Precuneus	L	0	-48	36	115	
Precentral Gyrus	L	-51	3	57	128	
Fusiform Gyrus	R	36	-57	-18	29	
Supramarginal Gyrus	R	54	-30	33	22	
Inferior Temporal Gyrus	L	-54	0	-33	21	

Table 16: Coordinates of peek activation from significant clusters

Derived from the interaction effect between task type and difficulty level (Figure 6)

		MNI Coordinates (mm)			Cluster Size
Peak Region	Hemisphere	x	y	z	(in voxels)
Fusiform Gyrus	R	33	-69	-6	390
FusiformGyrus	L	-30	-48	-12	151
Middle Occipital Gyrus	R	33	-87	21	154

Table 17: *F* test on ROIs derived from peak activations in first linear contrast (Figure 7)

Hemisphere	ROI	*F*	*p*	*Significance*
R	Angular Gyrus	27.4451648	1.09E-24	***
L	Angular Gyrus	58.830953	1.53E-48	***
L	Medial Frontal Gyrus	46.331621	1.2913E-39	***
R	Superior Frontal Gyrus	74.6265607	6.41572E-59	***
L	Superior Frontal Gyrus	53.655558	6.26545E-45	***
L	Posterior Cingulate	40.583448	2.87134E-35	***
L	Superior Temporal Gyrus	38.688192	8.45921E-34	***
L	Superior Temporal Gyrus	112.146061	2.40052E-80	***
R	Superior Parietal Lobe	29.0844905	4.61387E-26	***
L	Superior Parietal Lobe	53.210326	1.29764E-44	***
R	Middle Frontal Gyrus	110.925708	1.04108E-79	***
L	Middle Frontal Gyrus	37.663201	5.36534E-33	***
R	Insula	51.761685	1.40599E-43	***
L	Insula	40.3566285	4.29496E-35	***
R	Thalamus	49.7257084	4.15086E-42	***
L	Thalamus	35.598768	2.30189E-31	***
L	Middle Cigulate Cortex	50.678649	8.46669E-43	***
L	Precuneus	13.3135849	3.15556E-12	***
L	Precentral Gyrus	43.7387962	1.127E-37	***
R	Fusiform Gyrus	47.8487155	9.77588E-41	***
R	Supramarginal Gyrus	38.6794985	8.59234E-34	***
L	Inferior Temporal Gyrus	42.6429354	7.62109E-37	***
R	Fusiform Gyrus	50.643875	8.97091E-43	***
L	FusiformGyrus	52.15953	7.29228E-44	***
R	Middle Occipital Gyrus	12.272477	2.88688E-11	***

Table18: Coordinates taken from (Yeo et al., 2011)

		MNI Coordinates (mm)		
Peak Region	Hemisphere	x	y	z
Posterior Cingulate Cortex	L	-3	-49	25
Medial Prefrontal Cortex	L	-7	46	-2
Parahippocampal Complex	L	-25	-31	-20
Retrosplenial Cortex	L	-7	-50	7
MT+	L	-45	-72	3
Frontal Eye Fields	L	-26	-6	48
Intraparietal Sulcus	L	-40	-37	42
Inferior Anterior Prefrontal Cortex	L	-41	55	4
Dorsal Anterior prefrontal Cortex	L	-31	39	30

Table19: *F* test on the effect of difficulty for each ROI in the (Yeo et al., 2011) coordinates

ROI	*F*	*p*	*Significance*
Posterior Cingulate Cortex	22.185891	3.46646E-20	***
Medial Prefrontal Cortex	67.5930934	2.05265E-54	***
Parahippocampal Complex	0.390754	0.855237696	
Retrosplenial Cortex	24.0630932	8.20E-22	***
MT+	6.2054539	1.33108E-05	***
Frontal Eye Fields	35.7694798	1.68E-31	***
Intraparietal Sulcus	16.404732	4.75E-15	***
Inferior Anterior Prefrontal Cortex	14.5497511	2.31386E-13	***
Dorsal Anterior prefrontal Cortex	1.6017409	0.1578593	

Figures

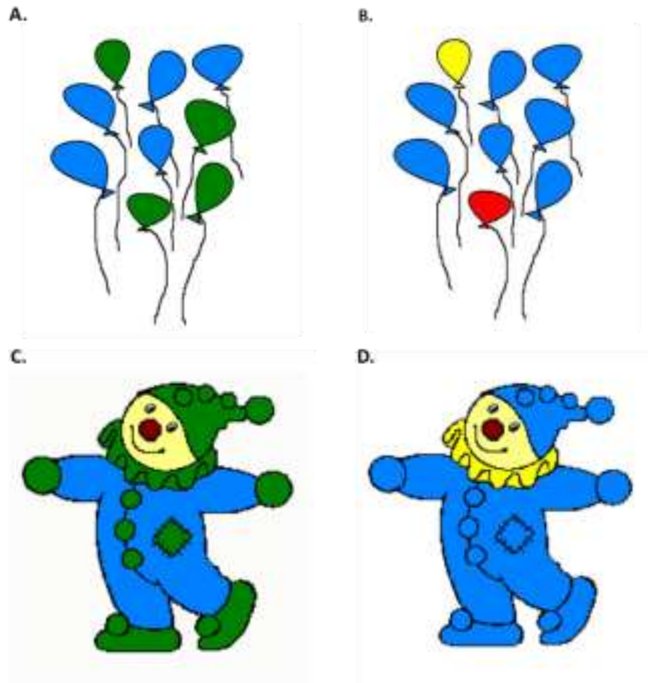


Figure 1: Baseline images and stimuli for the CMT. A) Balloon stimuli made up of all green and/or blue balloons were control stimuli because they do not include any colours of interest. Participants were required to press the "3" key to indicate that no comparison was being made. B) An example of difficulty level 3 where M-demand is n +1 = 3, where n was the number of colours of interest (i.e. yellow and red) and the blue balloons were to be ignored. C) The CMT clown presented participants with pictures of clowns where the face of the clown was to be ignored. The control stimuli for the CMT clown include all blue and/or green elements in addition to the clown face. Participants were required to press the "3" key to indicate that no comparison was being made. D) An example of difficulty level 3 where M-demand was n + 2 = 5, with one colour of interest (i.e. yellow).

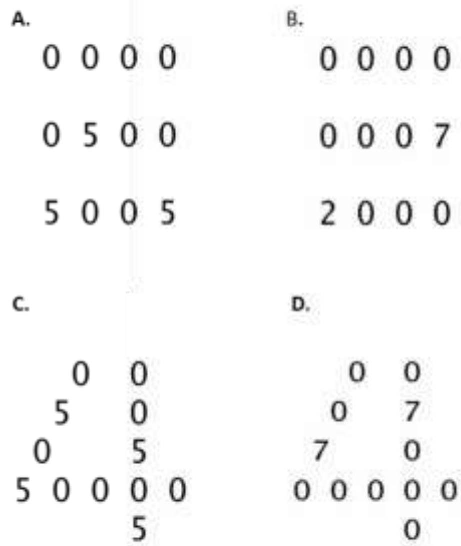


Figure 2: Baseline images and stimuli for the NMT. A) The NMT squares presented participants with a large rectangle made up of smaller numbers. Stimuli made up of all 0s and 5s were control stimuli because they did not include any numbers of interest. Participants were required to press the “3” key to indicate that no comparison was being made. B) An example of difficulty level 3 where M-demand was n +1 = 3, where n is the amount of numbers of interest (i.e. 2, 7). C) The NMT fours presented participants with images of a large 4 made up of smaller elements. The control stimuli for the NMT fours task included only 0s and/or 5s. D) An example of difficulty level 3 because there was one number of interest (i.e. 7) and M-demand was n + 2 = 3.

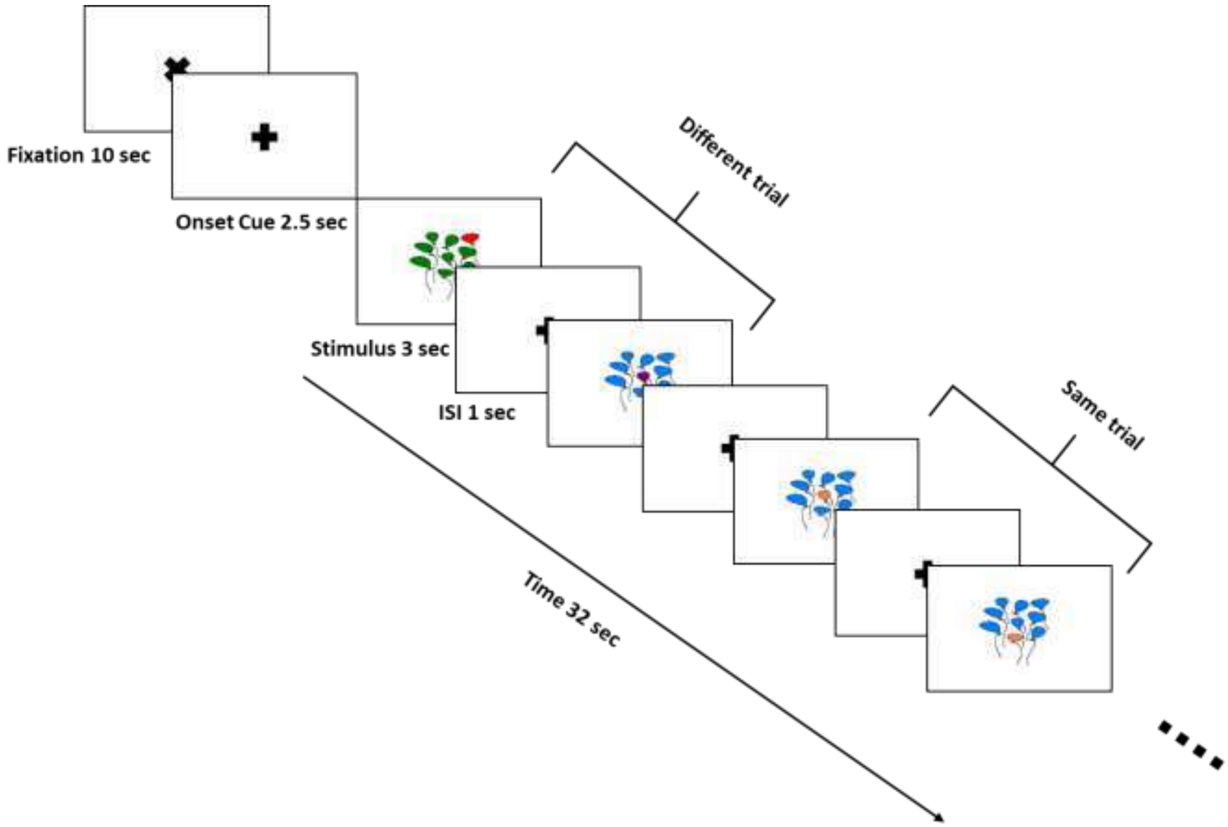


Figure 3: Task structure. The CMT and NMT presented the participant with a sequence of stimuli and required them to make a same/different judgement comparing a current stimulus to a previous one. Each block presented a series of eight images belonging to one difficulty level, determined by the number of elements that must be attended to. Each stimulus was presented for 3 s followed by an ISI of 1 s. Each block lasted 32 s, which included all eight images and a 1.5 s offset cue at the end of the block signifying that a period of fixation was to follow. Task blocks were interleaved by 10 s periods of fixation and a 2.5 s block-onset cue.

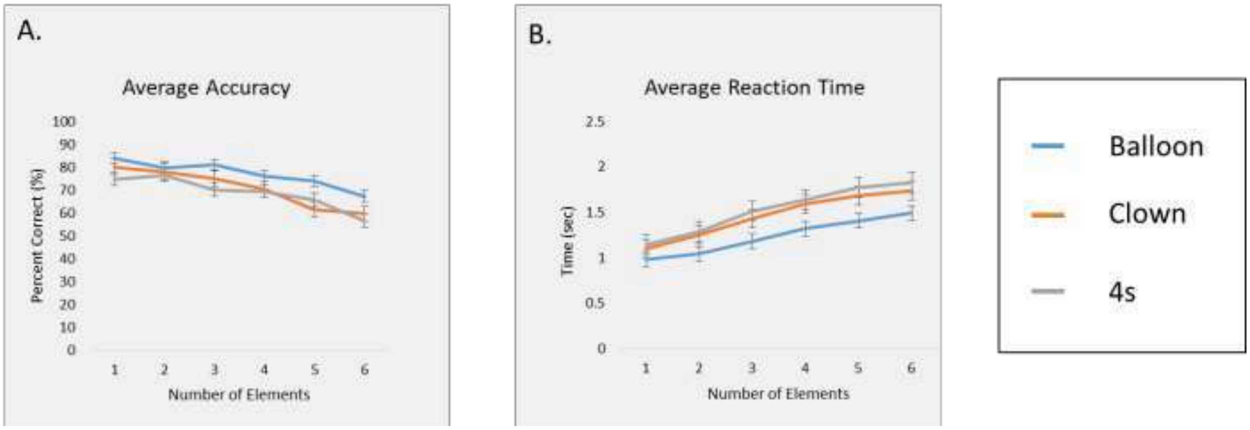


Figure 4: Accuracy and Reaction time. A) Average accuracy and B) reaction time across difficulty levels according to task type.

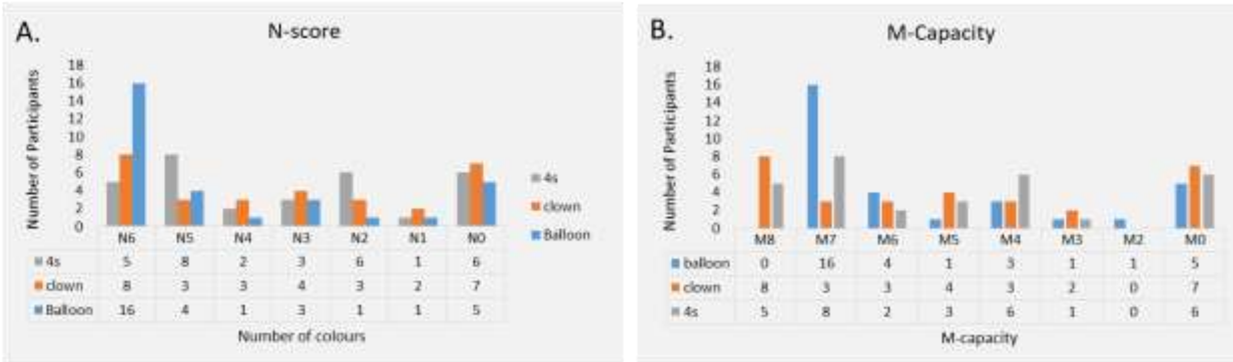


Figure 5: M-capacity. The highest level of difficulty that participants passed with over 70% accuracy in each task.

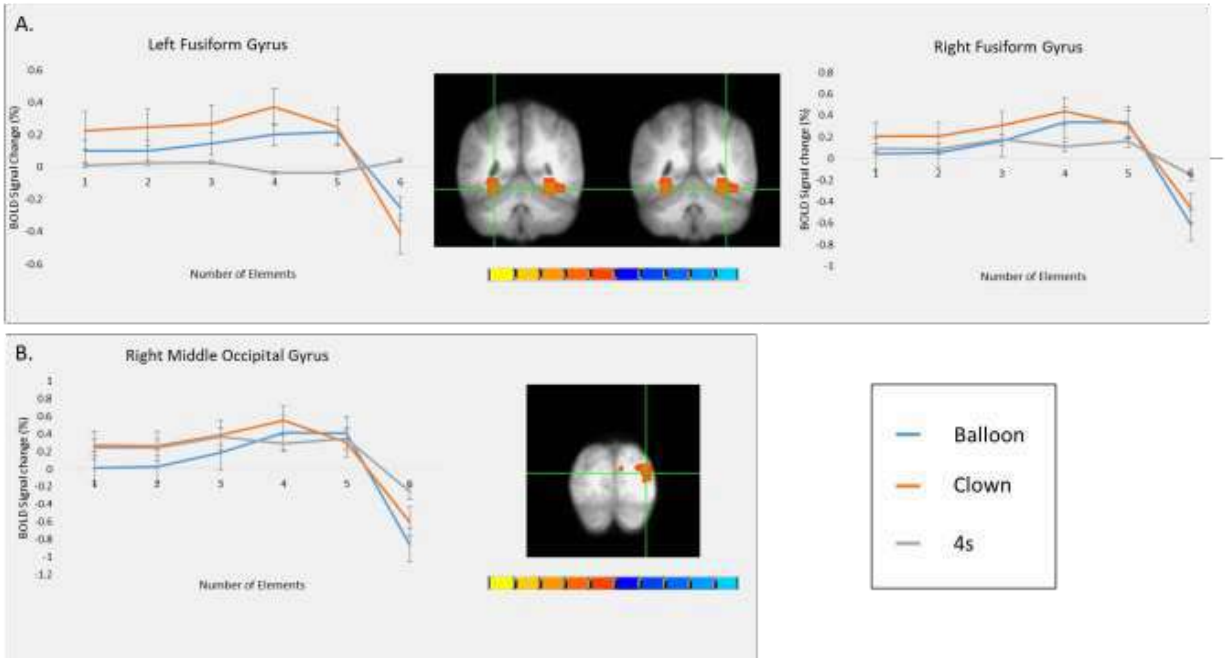


Figure 6: Interaction effect of task type and difficulty. The minimum cluster size was set to 20 voxels at q < .01 (FDR corrected). Average BOLD signal time course in each cluster plotted against difficulty level. Warm colours indicate regions showing a positive linear association between difficulty and activation, while the cool colours indicate regions showing a negative association.

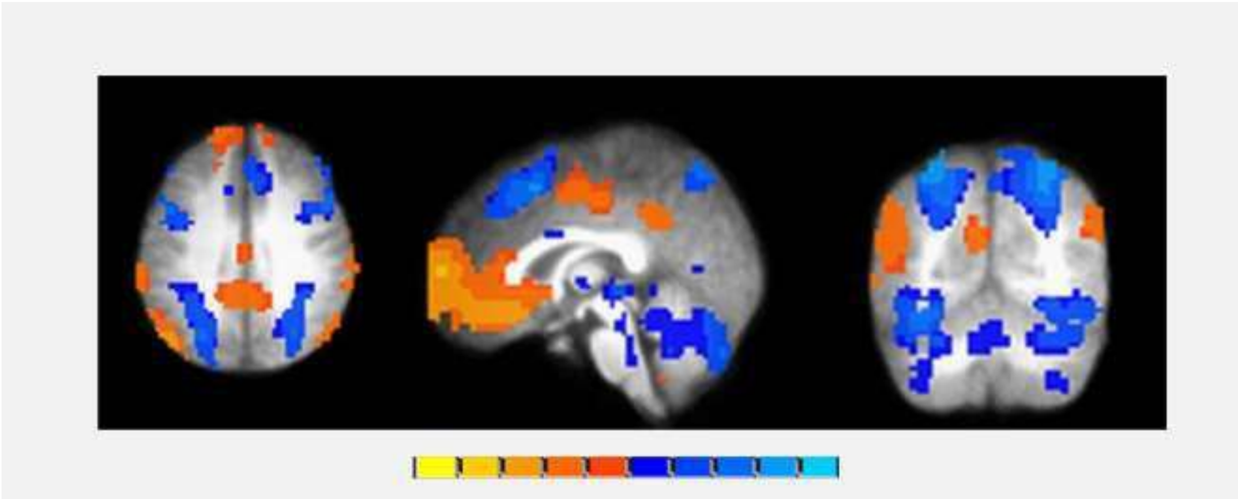


Figure 7. Linear contrast of difficulty levels one to six. Collapsing across tasks. Linear contrast identifying voxels that significantly associate to the pattern: D1<D2<D3<D4<D5<D6, conducted across all tasks, N= 23. The minimum cluster size was set to 20 voxels at q < .001 (FDR corrected). Warm colours indicate regions showing a positive linear association between difficulty and activation, while the cool colours indicate regions showing a negative association.

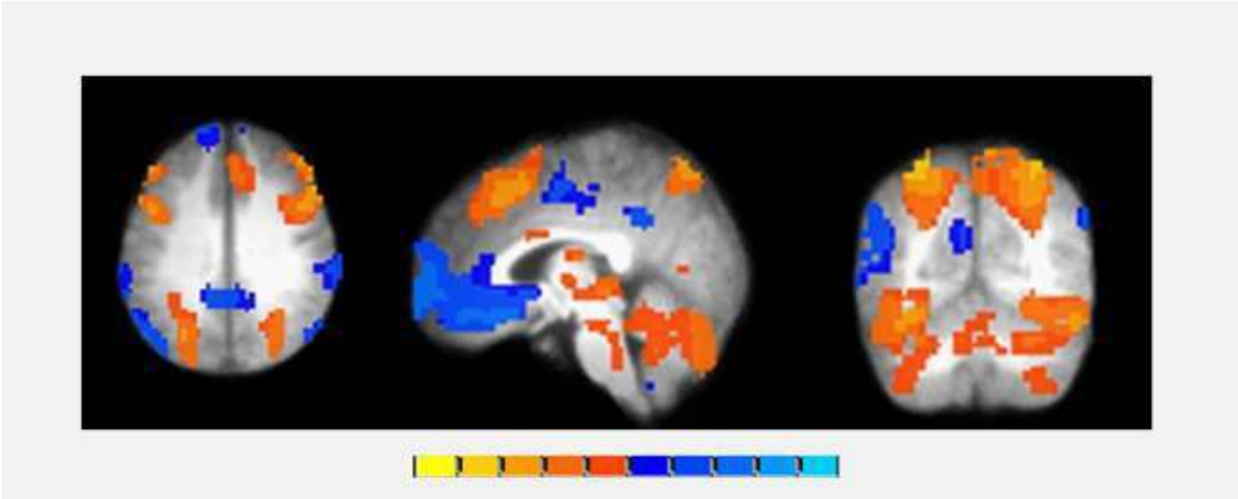


Figure 8: Quadratic contrast of all levels of difficulty. Collapsing across tasks. Quadratic contrast identifying voxels that significantly associate to the pattern: D1<D2<D3=D4>D5>D6, conducted across all tasks, N= 23. The minimum cluster size was set to 20 voxels at q < .001 (FDR corrected). Warm colours indicate regions showing a positive linear association between difficulty and activation, while the cool colours indicate regions showing a negative association.

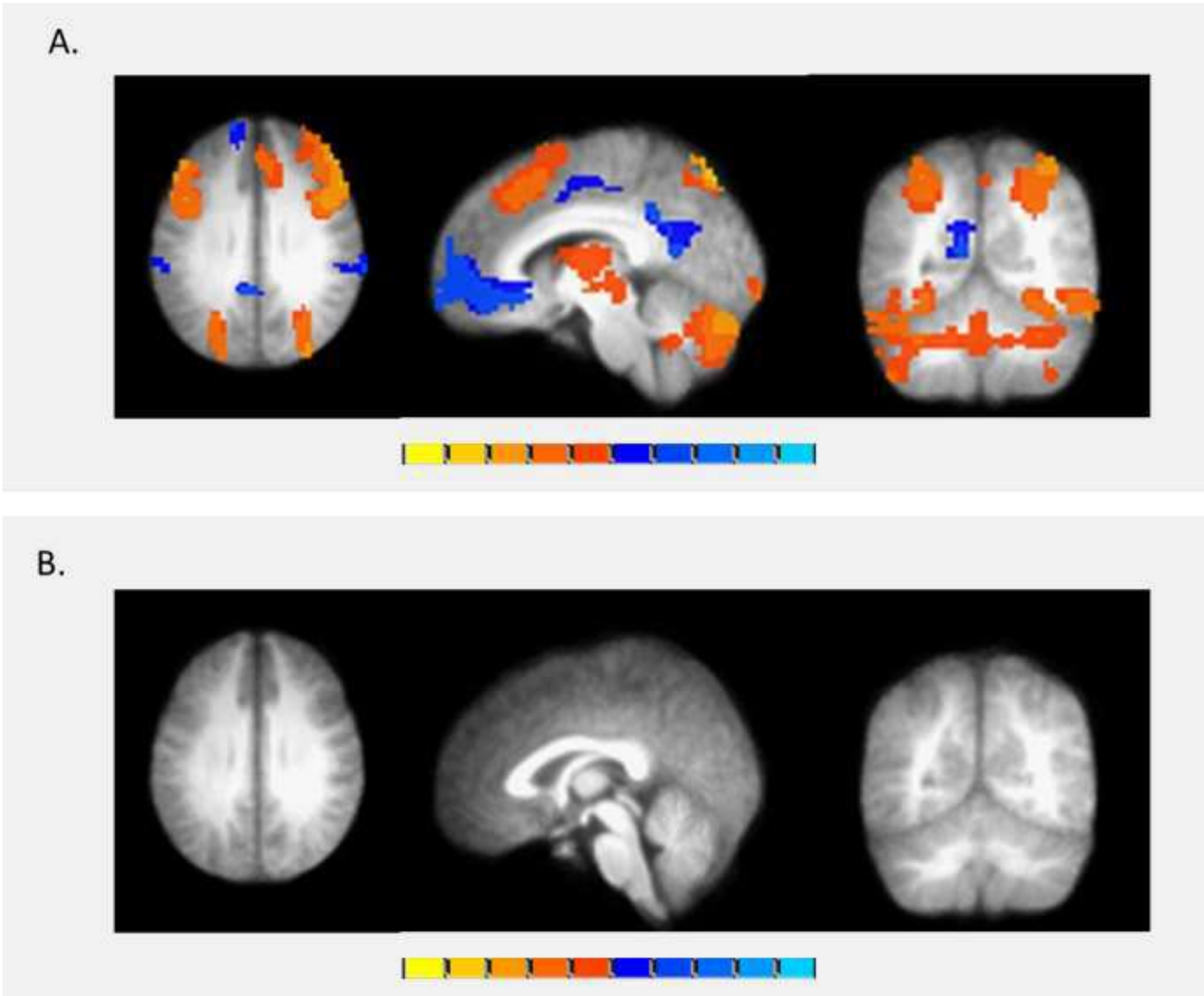


Figure 9: Linear and quadratic contrast of difficulty levels one to five. A) Linear contrast identifies voxels which significantly relate to the contrast: D1< D2< D3< D4< D5. The minimum cluster size was set to 20 voxels at q < .001 (FDR corrected). Warm colours indicate positive activation, while the cool colours indicate negative activation. B) Quadratic contrast identifying voxels that relate to a U shape function.

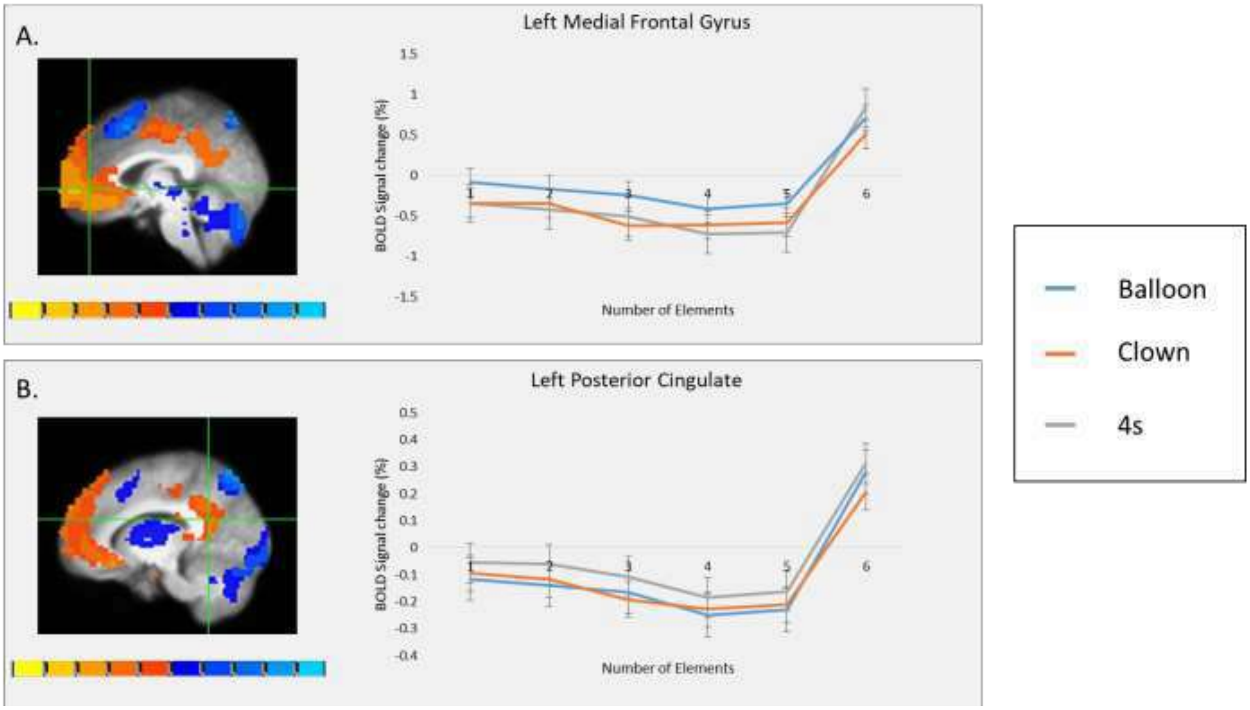


Figure 10: Default network ROIs, Arsalidou et al. (2013). The average BOLD signal time course extracted from ROIs derived from Arsalidou and Colleagues (2013). Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

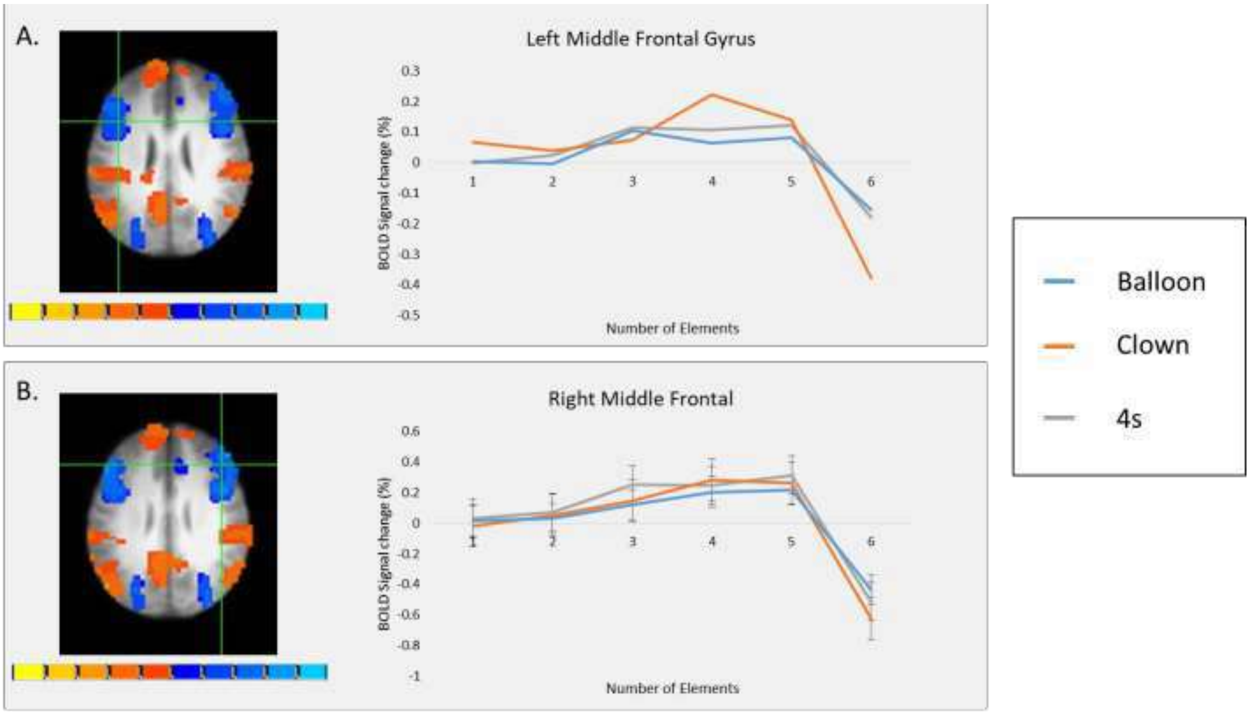


Figure 11: Frontoparietal control network ROIs, Arsalidou et al., (2013). The average BOLD signal time course extracted from ROIs derived from Arsalidou and colleagues (2013). Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

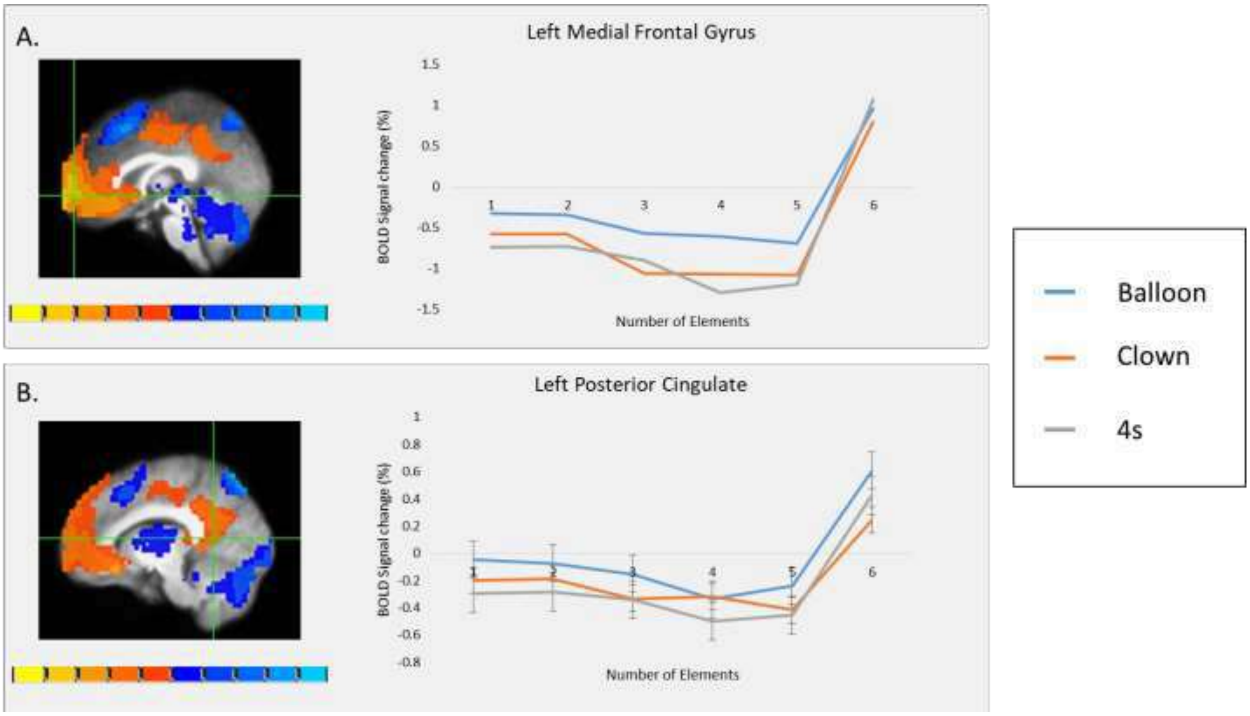


Figure 12: Default network ROIs, present investigation. The average BOLD signal time course extracted from ROIs derived from peak activations in the linear contrast shown in figure 7. Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

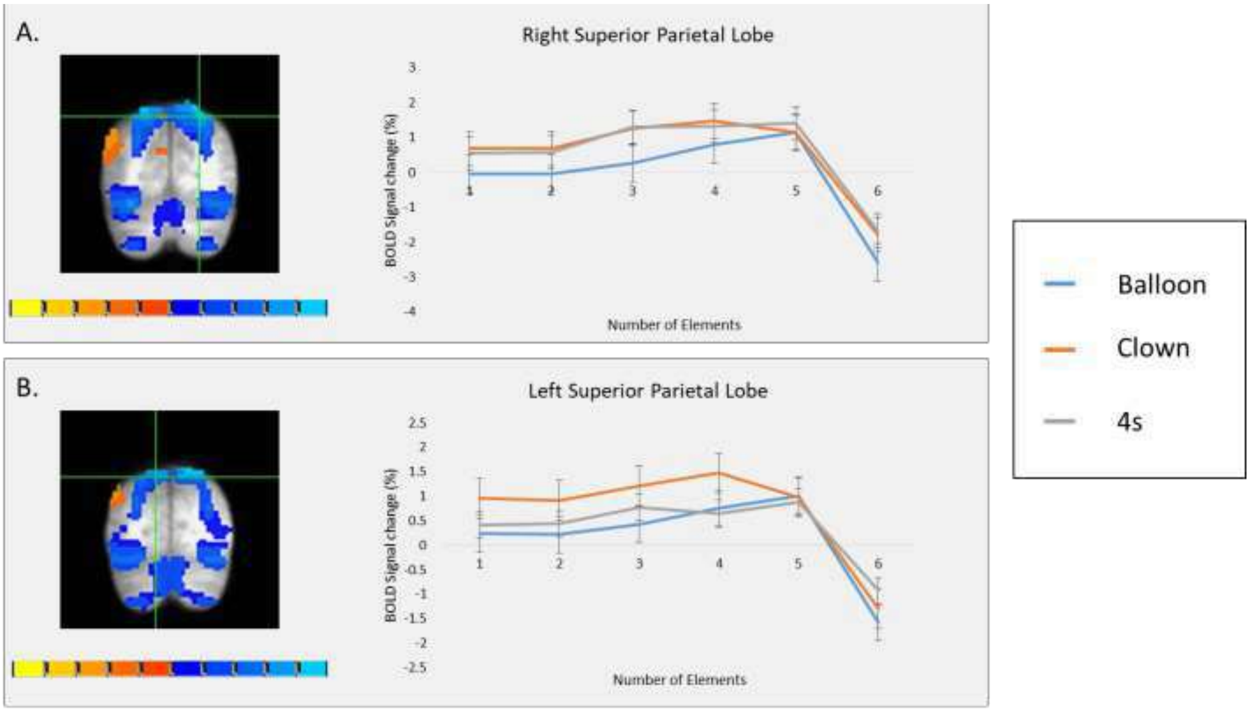


Figure 13: Dorsal attention network ROIs, present investigation. The average BOLD signal time course extracted from ROIs derived from peak activations in the linear contrast shown in figure7. Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

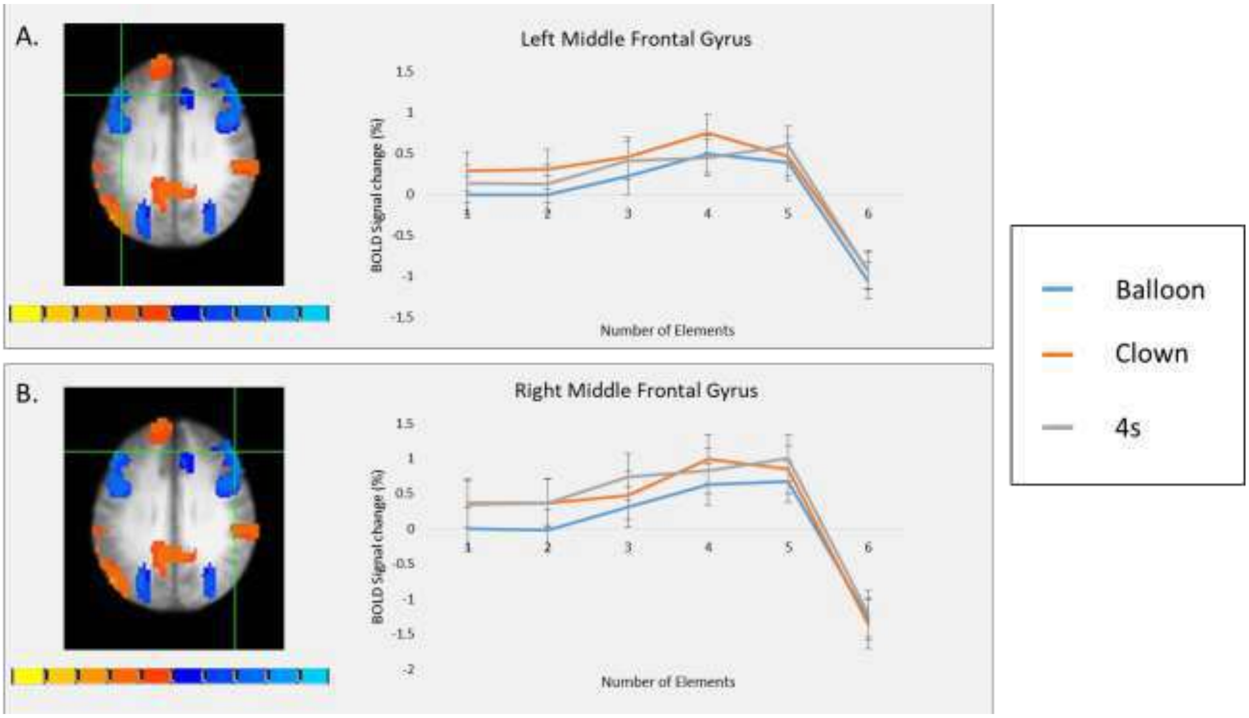


Figure 14: Frontoparietal control network ROIs, present investigation. The Average BOLD signal time course extracted from ROIs derived from peak activations in the linear contrast shown in figure7. Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

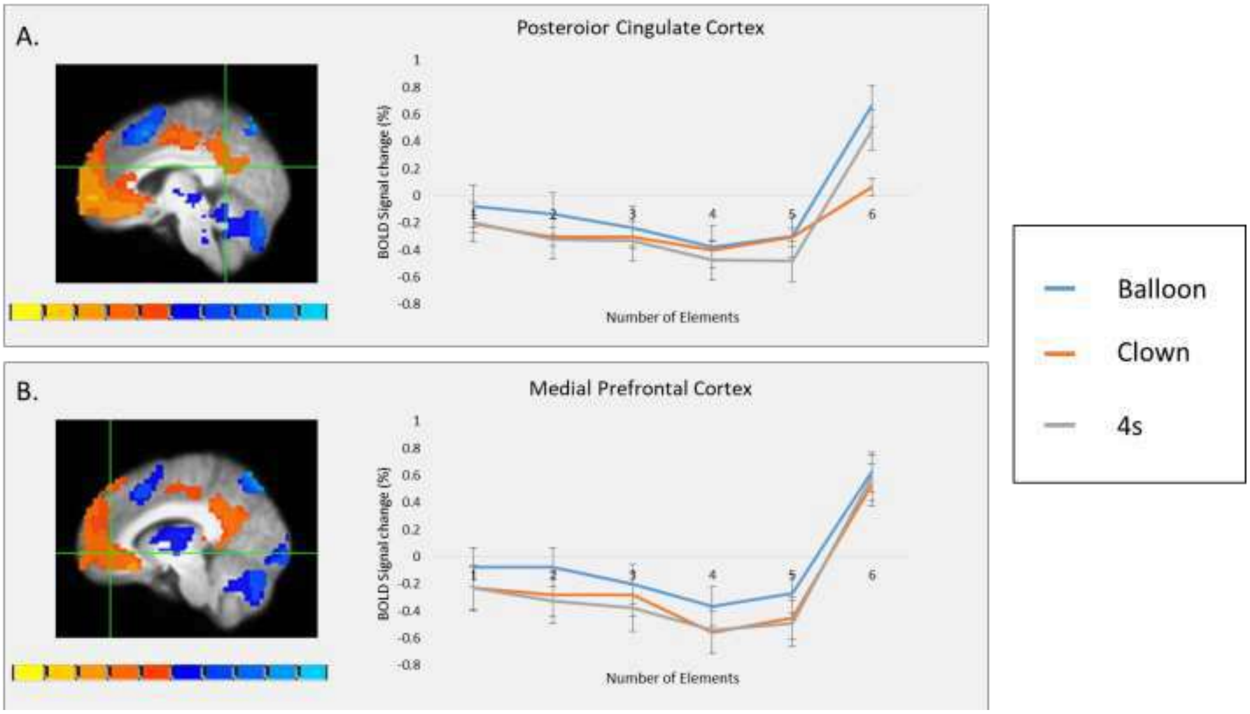


Figure 15: Default Network ROIs, Yeo et al., (2011). The Average BOLD signal time course extracted from ROIs derived from Yeo and colleagues (2011). Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

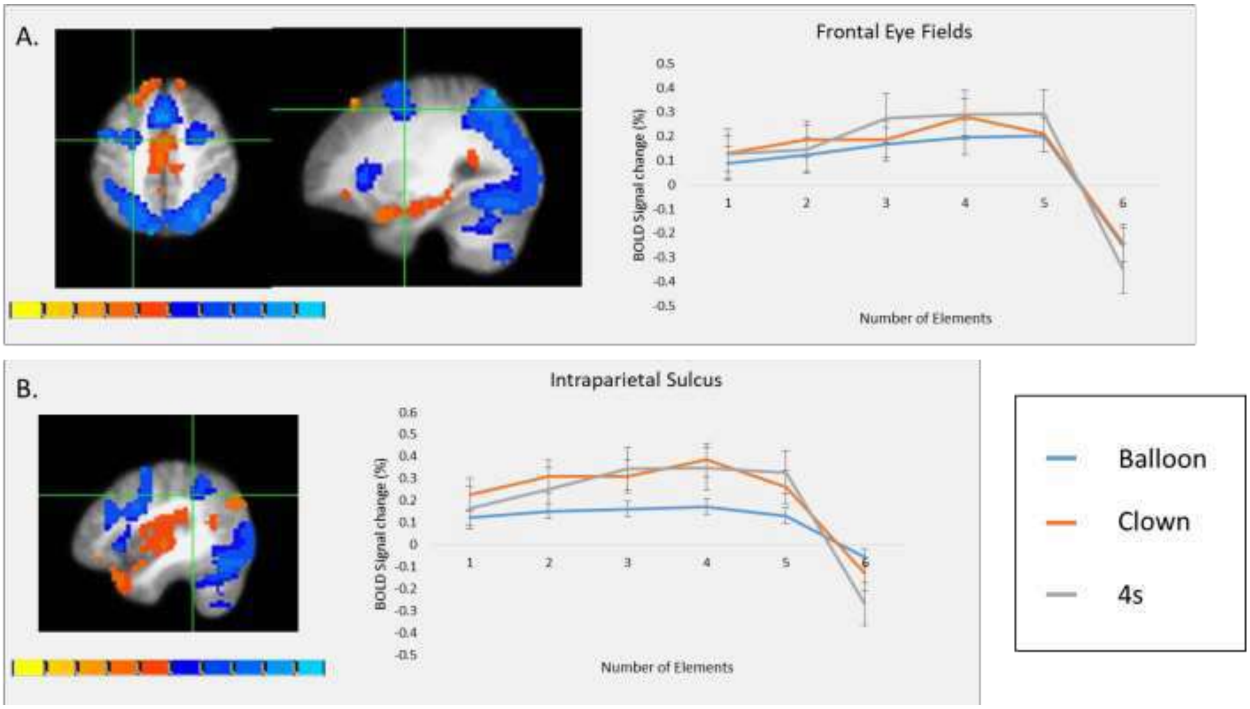


Figure 16: Dorsal attention network ROIs, Yeo et al., (2011). The Average BOLD signal time course extracted from ROIs derived from Yeo and colleagues (2011). Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.

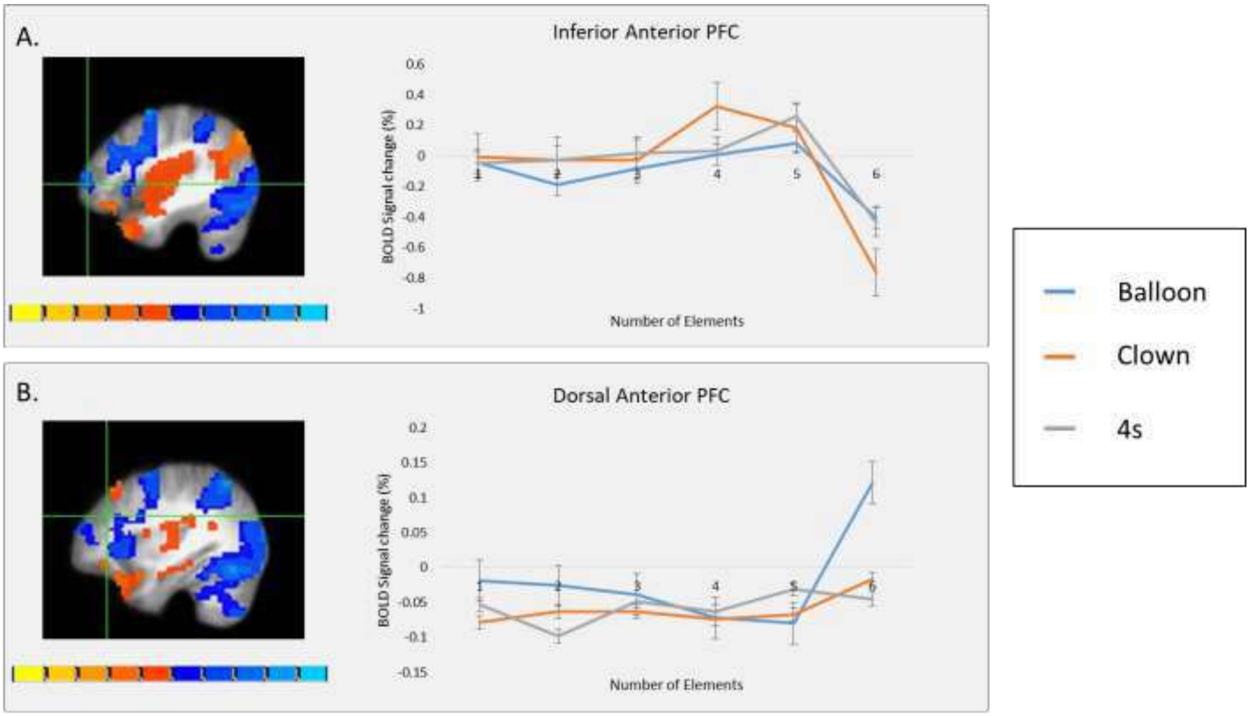
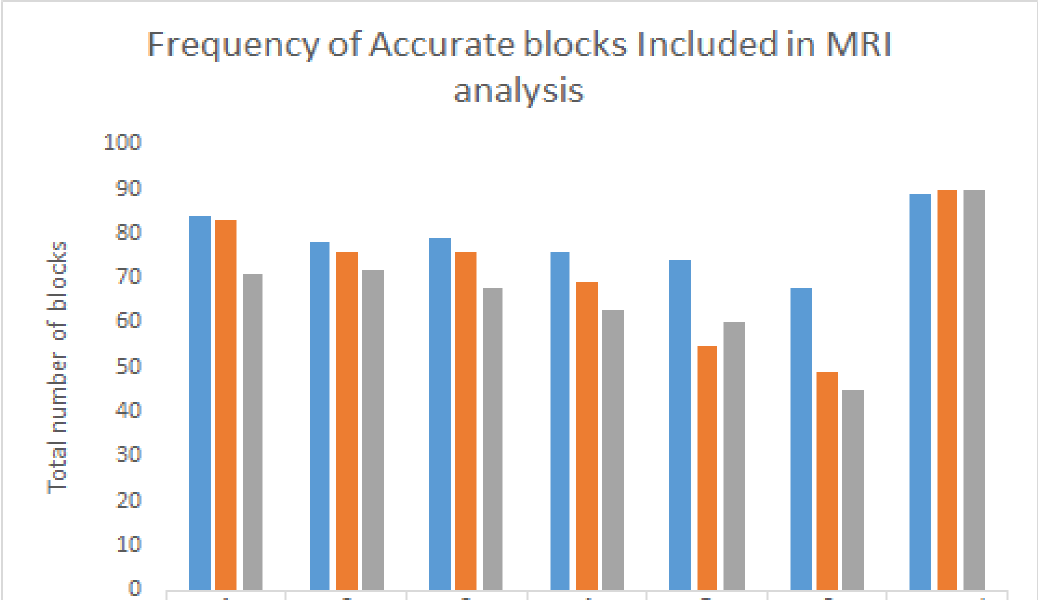


Figure 17: Frontoparietal control network ROIs, Yeo et al., (2011). The Average BOLD signal time course extracted from ROIs derived from Yeo and colleagues (2011). Activation map showing significant voxels from linear contrast in figure 7; warm colours indicate a positive linear association between difficulty and activation, while the cool colours indicate a negative association. Crosshairs indicate the location of the ROI coordinates.



	1	2	3	4	5	6	control
balloon	84	78	79	76	74	68	89
clown	83	76	76	69	55	49	90
4s	71	72	68	63	60	45	90

Figure 18: The Number of accurate blocks in each difficulty level for each task. Signifying the amount of blocks that were included in each condition for the fMRI group analysis.

Appendix A

Practice Session Instructions

Training Session – Explaining Task outside of Scanner with power point
This game is for both young children and adults. The instructions are designed to be clear to young children so please be patient with me.
Colors

- In this game there are a number of different colors.
- I will now show you the colors and I want you to name them. Ready?

Balloons

- In this game you are going to see a picture made up of balloons like this one.
- The balloons are going to be different colors.
- All colors are part of the game except blue and green.

Blue and Green balloons are not part of the game.
When you see blue and green balloons just look and respond with the 3rd button. When you see an X just look and do not respond
When you see a + that means a new picture is coming soon
RULES: Same/Different

- In this task you are going to look at pictures of balloons that have different colors.
- You must check whether the balloons have the same colors as the last time you saw them.
- Between pictures you will see a cross in the center of the screen. The cross is there to tell you where to look, so that you will be prepared when the next picture comes.
- When the cross comes you have to keep in mind the colors of the last picture.
- You will tell me whether/if the colors of the picture you are looking at are the same as in the last picture.
- Now I will show you 2 pictures of balloons and you must tell me if the colors are the Same or Different.

?T1
Same: Yes. The color was the same but in the second picture it was found in a different balloon. In this game it does not matter where the color is found.
Let's try another one. This is the first picture… Are the colors the same or different?
?T2
Yes, the colors are the same, even though they are found on different balloons.

What about this one? Are the colors the same or different?
?T3
The colors were the same. Remember that you must IGNORE blue and green. The other colors were still the same.
Let's try this one. Look at the pictures and tell me if the colors are the same or different.
?T4
The colors were Different because the color in the last one was not the same as the one before that.
What about this one?

?T5
The colors were Different between the two pictures.
Let’s get some practice (give positive feedback when participant responds correctly)
In the real game the pictures are going to appear one at a time. For responding in the Simulator and in the MRI scanner you will need 3 buttons, 1 for same, 2 for different and 3 for ignore.
To get a better idea on how the task works now I will show you some pictures on the same screen. So this is the first one,
For the first picture, you also need to press the 3rd button because we are not comparing it to anything else. Then the next appears. Do the pictures have the same or different colors as the last picture? But remember that blue or green do not count.
Same or different … how about now? … same or different …
Remember when you see balloons that are only blue and or green you look at them and press the 3 key. After that, you press the 3 key on the first new picture and start checking the colors again.
This is the first picture … then same or different.
**repeat same instructions only with the numbers for the squares and 4s task

Scripted Motion Reduction Training Instructions (participant is inside simulator)
Please keep your eyes fixated on the (+) in the center of the screen.

Even when you feel you are staying still, the tracker is still sensing little movements, generally within the green circle you see on the screen.
We are going to do a few example movements to demonstrate how still you need to be in the MRI scanner.

1. Take your finger and try to touch your nose
2. Take your foot and try to touch your opposite ankle
3. Say the word “hello” 5 times
4. Make a pretend cough
5. Make a pretend sneeze
6. Try to lift your lower back off of the platform and re-position yourself.

As you can see many of these movements cause the tracker to move around and in some cases even move beyond the green circle.
Now we are going to play a quick video and your goal is to stay as still as possible while the video is on. If the tracker senses movement beyond the green circle, the video will pause. Once the tracker moves back into the green circle the video will begin again.
Are you ready?
There are many movements that cause head motion even when you “feel” like your head is staying still. Because motion affects the quality of the images it is important to keep this in mind when you are in the real MRI scanner.

Scripted Practice Task Instructions (Participant is inside the simulator)
CMT – Balloon, Clown
Remember the rules of the Game!

- You have to check if the image has the same colors as the last time you saw it.
- The pictures are going to be on the screen for a very short time. You have to respond as quickly and as well as you can before the cross comes up.
- After you see 8 seconds of an (x) will come up telling you that you have a few seconds where you need to just keep your eyes open and pay attention. When the (+) comes up it means "get ready because a new image is about to come up".
- You must respond with the (1) button if the answer is "same" or (2) button if answer is "different". If the image is all blue and green then you need to respond with the (3) button for "ignore".
- If you think you made a mistake don't stop, just keep going.

Do you have any questions?

NMT – 4s
Remember the rules of the Game!

- You have to check if the image has the same numbers as the last time you saw it.
- The pictures are going to be on the screen for a very short time. You have to respond as quickly and as well as you can before the cross comes up.
- After you see 8 seconds of an (x) will come up telling you that you have a few seconds where you need to just keep your eyes open and pay attention. When the (+) comes up it means "get ready because a new image is about to come up".
- You must respond with the (1) button if the answer is "same" or (2) button if answer is "different". If the image is all 0s and 5s then you need to respond with the (3) button for "ignore".
- If you think you made a mistake don't stop, just keep going.

Do you have any questions?
NMT – squares
Remember the rules of the Game!

- You have to check if the image has the same numbers as the last time you saw it.
- The pictures are going to be on the screen for a very short time. You have to respond as quickly and as well as you can before the cross comes up.
- After you see 8 seconds of an (x) will come up telling you that you have a few seconds where you need to just keep your eyes open and pay attention. When the (+) comes up it means "get ready because a new image is about to come up".

- You must respond with the (1) button if the answer is "same" or (2) button if answer is "different". If the image is all 0s and 5s then you need to respond with the (3) button for "ignore".
- If you think you made a mistake don't stop, just keep going.
- This task is going to be longer than the ones we did so far, we have 4 runs that are 5.5 mins each.

Do you have any questions?
Countdown runs
Great job! We have (3, 2, 1) runs left for this task!

Scripted Briefing of Participant for MRI session
In the real MRI
-the screen is not as bright as this one in the simulator
-there are more cushions available so you will likely feel even more comfortable than you are here
-the MRI makes very loud noises, recommend going on youtube and seeing an example of what a real scanner sounds like.
-the real MRI is the same size on the inside but the machine around it is much much bigger!
-you will be doing 3 tasks in the scanner (balloon, clown, 4s). Each one is the same length and structure as the squares task you did today.
-We start with a rest run (10 mins of you fixating on an X, try not to fall asleep!), then the balloons, then clown, then a structural scan where you can close your eyes and rest, then the 4s.

Appendix B

MRI Session Instructions

Scanning Day instructions
Task Scan:
Remember the rules of the Colour tasks and the Number tasks

- You have to check if the image has the same numbers/colours as the last time you saw it.
- The pictures are going to be on the screen for a very short time. You have to respond as quickly and as well as you can before the cross comes up.
- Remember to ignore blue and green for the colours and 0s and 5s for the numbers.
- After you see 8 seconds an (x) will come up telling you that you have a few seconds where you need to just keep your eyes open and pay attention. When the (+) comes up it means "get ready because a new image is about to come up".
- You must respond with the (1) button if the answer is "same" or (2) button if answer is "different". If the image is all 0s and 5s then you need to respond with the (3) button for "ignore".
- If you think you made a mistake don't stop, just keep going.

Do you have any questions?