# Picture this: an investigation of the neural and behavioural correlates of mental imagery in childhood and adulthood with implications for children with ADHD

Kathryn Elizabeth Bates

University College London

May 2021

A thesis submitted for the degree
of Doctor of Philosophy

I, Kathryn Elizabeth Bates confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Kathryn Elizabeth Bates

# Abstract

Mental imagery (MI), a vital tool in supporting memory and learning, is defined as the ability to generate and manipulate mental images in mind in the absence of sensory input. Despite its importance, there is limited understanding of the development of MI, or the developmental relationship between MI and visual working memory (VWM). In adults, it is speculated that individual differences in VWM capacity depend on variable recruitment of MI strategies. However, this has not been tested directly. The aims of this thesis are to address these gaps in the literature.

With respect to development, findings demonstrate that MI is visually depictive in nature in children from age 6 (in support of depictive theory of MI) and that MI is a multi-faceted function rather than a unitary construct. That is, components of MI (image generation, image maintenance, mental rotation, image scanning) develop separably from 6 to 11 years, although note that image maintenance and mental rotation become related in adulthood.

No relationship was found between components of MI and VWM in adulthood, typically developing children or in children with ADHD (age 8-14 years). Contrary to predictions, the ADHD group demonstrated broadly typical performance in each MI component and in VWM and no syndrome-specific profile of MI/VWM performance. This suggests that MI is not a weakness in ADHD and could be a useful learning tool for this group.

Exactly how individual differences in MI support VWM was tested in adults. Findings demonstrated that adults exert wilful control over the visual precision and capacity of visual representations within VWM. However, individual differences in both subjective MI vividness and quantity ratings did not map onto the neural correlates of VWM (contralateral delay activity, anterior directing attention negativity). Thus, it is concluded that the subjective experiences of MI are distinct from implicit visual representations.

## Impact statement

The impact of the findings from this thesis firstly pertain to current theoretical frameworks of the development of MI, the relationship between MI and VWM, and how MI might support memory and learning. First, the evidence presented here suggests that MI is a multi-faceted function, which develops as separable components throughout childhood. The novel battery of tasks introduced in this thesis provide a framework for examining individual differences in the components of MI from age 6 years in both typically and atypically developing children. The findings presented here are somewhat contrary to the current narrative in the literature in that the evidence suggests components of MI and VWM are unrelated. It has been recently argued that individual differences observed in VWM capacity are likely dependent on the recruitment of MI (i.e., visual strategies) in VWM (Pearson & Keogh, 2019). However, the findings presented here demonstrate that individual differences in subjective ratings of MI (a common measure of the sensory experience of MI) are disconnected from the visual precision (the precision at which visual information is held in mind) and capacity (the number of visual items held in mind) of visual representations in VWM. This contests the recent view that MI and VWM are synonymous and instead raises an important distinction between the sensory experience of MI and implicit visual representations recruited to support memory. This has implications for how MI is measured and conceptualised with respect to its contribution to learning and memory moving forward.

This battery of MI tasks has since been applied to investigate how components of MI contribute to mathematical calculation in children (Bates et al., 2021). The findings suggested that mental rotation is specifically important for mathematical calculation, as opposed to more visual components of MI. Such findings have implications with respect to the types of strategies that might be encouraged in mathematical calculation learning. The MI battery could therefore be applied to future research examining the variable contribution of the components of MI to a range of mathematical skills and beyond to other academic outcomes. Moreover, studies in typical development have found a causal link between mental rotation and mathematical calculation skills (Gilligan et al., 2019; Mix et al., 2020). In this thesis it was found that children with ADHD present with typical levels of abilities in components of MI, including mental rotation. Therefore, these results imply it might be possible to compensate for the negative impact of VWM impairments in children with ADHD by improving mathematical skills by training mental rotation.

# Dissemination

## Peer reviewed publications

*Data from Chapter 2 are published as a Journal Article in Cognitive Development:*

**Bates, K. E.**, Farran, E. K. (2021). Mental imagery and visual working memory abilities appear to be unrelated in childhood: evidence for individual differences in strategy use. *Cognitive Development. 60*, 10112. https://doi.org/10.1016/j.cogdev.2021.101120

*Collaboration from this thesis using data from Chapter 2 are published as a Journal Article in Mind, Brain and Education:*

**Bates, K. E.,** Gilligan-Lee, K., & Farran, E. K. (2021). Reimagining Mathematics: The Role of Mental Imagery in Explaining Mathematical Calculation Skills in Childhood. *Mind, Brain, and Education*. https://doi.org/10.1111/ mbe.12281

## Conference proceedings

*Data from Chapter 4 were presented at the Cognitive Neuroscience Society online conference:*

**Bates, K. E.** , Smith, M. S., Farran, E. K., Machizawa, M. G. (2021), How visual is visual working memory? Investigating the metacognitive link between visual imagery and visual working memory. Poster presentation.

*Data from Chapter 3 were presented at the European Society for Cognitive Psychology Conference, Tenerife, Spain:*

**Bates, K. E.**, Farran, E. K. (2019), Mental Imagery in typical development and ADHD. Poster presentation.

*The pilot data and methods outline from Chapter 2 were presented at the British Psychological Society Developmental Section Annual Conference, Liverpool, UK:*

**Bates, K. E.**, Farran, E. K. (2018), Testing a developmentally appropriate mental imagery task battery. Poster presentation.

## Acknowledgements

I would first and foremost like to thank my supervisors, Prof. Emily Farran and Dr Marie Smith, for your guidance, wisdom, and kindness throughout this process. A huge thank you goes to Emily for pushing me to explore my ideas and providing expertise to help formulate the research, and for always having complete faith in my abilities as a researcher. You have gone over and above to support me and my work; the PhD is independent in nature, but I have never felt alone thanks to your open door.

I would like to thank everyone who participated in my research. A special thank you to all the staff, parents, and children at the schools I worked with for welcoming me into your classrooms with such enthusiasm and flexibility. Thank you to all the parents and children who gave up their time on weekends and school holidays to participate. This thesis would not have been possible without you.

I have been very lucky to have met many wonderful people who have become collaborators, mentors and friends and I would like to thank you all for your contribution to my growth as a researcher and to the quality of this thesis. I'd firstly like to thank my colleagues in the CoGDeV Lab for providing such a nurturing and stimulating environment. I feel very privileged to have completed my PhD in this group. Thank you to my BMA friends; especially, Dr Emily Thomas and Dr Gurmukh Panesar. Thank you for the long breaks in the garden, the delicious home-cooked treats, and the research- and non-research chats. It has been a pleasure to share this journey with you. I'd also like to thank the Women in Working Memory writing group hosted online throughout this difficult year by Dr Alicia Forsberg and Dr Kirsten Adam. I have looked forward to joining the sessions each week and have thoroughly enjoyed writing with such an inspirational group of women. My thanks go to Dr Pawel Matusz; you have been a voice of reason and such a selfless source of support and kindness. I am very grateful for your mentorship over the past year and a half. Finally, a special thank you goes to Dr Maro Machizawa. Thank you for always responding to my lengthy emails with such patience and clarity, and for facilitating meetings and interesting discussions across time zones. I have learnt so much from working with you; it has been one of the highlights of my PhD.

Thank you to my family for the love and support not only during my PhD, but always. I'd especially like to thank my parents: Carol and John Elderfield and Julie and Steve Bates. Carol and John, thank you for always being there, being proud of me no matter what and making all this possible. Julie and Steve, thank you for always encouraging me, whisking me away on adventures and teaching me to work hard and play harder. Special thanks go to my big, little sister, my best friend. Your friendship and love mean the world to me and the frequent FaceTimes and little gifts in the post have picked me up when I've needed it most. My thanks also go to Jane and Adrian Exell. Thank you, Jane for supporting my data collection in your school; you made this experience so enjoyable. Your passion is admirable and has certainly rubbed off on me. Thank you to both Jane and Adrian for welcoming me into your home during data collection; your unbound enthusiasm is infectious, and your support has been such an important part of this journey.

I would like to thank my friends for their patience and for dragging me away from my laptop for much-needed breaks. I would especially like to thank Mallory Durran; the smartest and most caring woman I know. Your zeal for research is so inspiring. A special thank you goes to Hannah Rea and Christopher Bounds. Chris, thank you for always being interested in my work and for asking me thought-provoking questions. Hannah, I could not wish for a more thoughtful, kind, and hilarious friend. Our friendship has carried me through to the end of this process.

To my partner Josh Exell. You have been there for the good times, the sad times, the stressful times; and always with the kindest smile on your face. You have taught me to make the most of every day, and even when I've struggled to believe in myself, your belief in me has never faltered. Your love, compassion and energy has made this the most incredible experience - onto the next adventure!

# Table of contents

## List of figures

## List of tables

# Chapter 1: Introduction to thesis

## 1.1. Introduction

*“There are many differences between humans and the rest of the species on earth, but one that has been expressed is that we alone are able to imagine the future.”*

- *David Attenborough, A Life On Our Planet*

Our ability to generate perceptual phenomena in mind allows us to contemplate the future and remember the past, whilst navigating through the present. Mental imagery (MI) is defined as the ability to generate visual mental images in mind in the absence of sensory input (Kosslyn, 1980). As an aid to memory and learning, MI has been implicated in reading comprehension and language development, as well as problem-solving in mathematics (Bizzaro et al., 2018; Commodari et al., 2020; Guarnera et al., 2019; Hegarty & Kozhevnikov, 1999; Sadoski & Quast, 1990; Sadoski & Paivio, 2004). Thus, MI is a vital problem-solving tool for an individual to make sense of their world and to excel in learning. MI is consistently likened to visual working memory (VWM); the ability to maintain and manipulate visual information in mind (Baddeley, 2003; Baddeley & Andrade, 2000; Cowan, 2001; Logie, 1995). Not only is there overlap between the theoretical definitions of MI and VWM, but research with adults has also demonstrated evidence for shared neural (Albers et al., 2013) and behavioural mechanisms (Keogh & Pearson, 2011, 2014) between MI and VWM. Consequently, it has recently been conceptualised that MI ability might underpin individual differences in VWM ability (Pearson & Koegh, 2019); however, this evidence is limited to adults. Given what we know about how MI contributes to learning, problem-solving and memory, it is surprising that research is yet to investigate how MI presents in children with impairments in working memory, specifically, children with attention-deficit/hyperactivity disorder (ADHD). Working memory deficits in this group are linked to poorer academic outcomes (Friedman et al., 2018; Orban et al., 2018; Sjöwall et al., 2017), however training visual working memory does not lead to lasting improvements in working memory, or transfer to academic outcomes, in children with ADHD (Cortese et al., 2015; Rapport et al., 2013). Clearly, research is required to establish how MI is related to visual working memory in both typically developing children and children with ADHD. The overarching aim of this thesis is to characterise how MI abilities develop throughout childhood (**Chapter 2**) and how MI presents in children with ADHD (**Chapter 3**), as well as to examine how individual differences in MI support VWM in typically developing children, children with ADHD and adults (**Chapters 2, 3 and 4**). This will provide a novel contribution to the field by delineating the relationship between MI, VWM and attention throughout typical and atypical development and in adulthood, which in turn has implications for understanding how MI contributes to learning and problem-solving.

## 1.2. MI as a format of thinking vs. MI as a cognitive function

There are two key strands to the research that have investigated the cognitive mechanisms of MI. The first has focused on the sensory experience of MI and aimed to determine the format of mental images. This research has examined the relationship between MI and visual perception to determine how visual representations are constructed in the brain and has investigated the role of MI in visual working memory (VWM) to establish the origin of individual differences in VWM capacity. The second key avenue of research has focused on the computational model of MI. This body of research has provided evidence for separable components in adulthood. This evidence suggests that MI is made up of distinct sub-abilities, but also that MI can be differentiated from verbal skills and reasoning abilities (verbal and non-verbal) (Kosslyn, 1994; Kosslyn et al., 1984).

This formed two defining attributes of the early conceptualisation of MI; first, that representations generated in mind are distinct from symbolic and abstract representations of general thought (Kosslyn, 1996) and second, the identification of a sub-component model of MI was imperative in conceiving that MI is not a single, unitary function, but is in fact multi-faceted, whereby individual differences in varying sub-components are present (Kosslyn, 1980).

Despite its importance, the latter has been largely neglected in recent decades. The focus of the research conducted over the last four decades has endeavoured to decipher the format of representations generated in MI, alongside the correlated but distinct functions of visual perception and VWM. Due to the private, unobservable nature of MI, research has primarily adopted neuroimaging methods to examine the format of MI. In turn, research using behavioural paradigms to investigate the multi-faceted cognitive function of MI is limited. Therefore, to provide a complete picture of the evidence for the mechanisms of MI to date, I will first review neural evidence for the format of MI followed by a review of the predominantly behavioural research examining the cognitive function of MI. Ultimately, I will put forward the argument that in order to fully characterise MI in development and adulthood, it is necessary to adopt a neurocognitive approach to investigate how MI develops as a multi-faceted construct (**Chapters 2 and 3**) and to examine how individual differences in the format of MI interacts with VWM (**Chapters 2, 3 and 4)**.

In **Section 1.2.1**. I will review the neural evidence for visually depictive representations in MI, making the argument that a) this body of research has shown that MI is not merely a visual representation emulating visual perception, and b) individual differences in MI contribute to behavioural attributes and neural correlates of VWM. In the following section, **Section 1.2.2.**, I will review cognitive evidence for a separable-component model of MI in child and adult populations. Here, I will formulate the argument that to fully examine MI abilities in childhood and adulthood and to consider their role in learning, it is necessary to examine MI as a multi-faceted construct and adopt a neurocognitive approach to investigating individual differences in MI throughout childhood and adulthood. This is the approach adopted in this thesis.

### 1.2.1. Deciphering the format of MI: shared neural mechanisms with visual perception and visual working memory

Decades of neuroimaging research on the format of mental images has provided extensive support for a depictive theory of MI. The depictive theory of MI proposes that mental images can involve percept-like, visual representations that are dependent on functional space in the early visual cortex, i.e., spatial relations map onto corresponding coordinates in topographically organised areas of the cortex, and visual perceptual properties are made accessible and explicit (Ganis, 2013; Kosslyn et al., 2006). This view was originally disputed by the propositional theory, which argues that mental representations are only depicted as symbolic, language-like descriptions, thus rendering the use of depictive mental images redundant (Pylyshyn, 1981, 2002). However, given the overwhelming evidence for a depictive theory of MI, it is now recognised that while information can be represented and stored in a variety of formats, visually depictive, percept-like representations can be recruited during MI and therefore a depictive theory of MI is supported (Pearson & Kosslyn, 2015). Thus, the remainder of this section focuses on the depictive theory only.

Evidence for a depictive theory of MI is largely derived from research examining the relationship between MI and visual perception with regards to shared neural mechanisms. A prominent framework of visual processing conceives of the dorsal and ventral visual pathways (Goodale & Milner, 1992; Mishkin et al., 1983; Ungerleider & Haxby, 1994).The dorsal visual stream, conceptually termed as the “where” pathway, is an occipitoparietal network between the early visual areas and prefrontal cortices via parietal lobules, that is involved in spatial perception. On the other hand, the ventral visual stream, the “what” pathway, denotes a hierarchical visual system stemming from the early visual cortices via an occipitotemporal network to specialised subcortical structures, including medial temporal regions and the hippocampus. It would be reductionist to assume that the visual processing pathways are wholly spatial vs. visual in that visual features nearly always entail spatial dimensions, and that there is evidence for neural connections between the two pathways (Van Essen et al., 1992). However, this account has provided a robust framework for understanding the processes of visual perception and object recognition within the ventral visual stream (Kravitz et al., 2011, 2013). Research investigating selective activation throughout the ventral visual stream has been central to our understanding of how stimulus content is represented in MI. In the next subsections I will review evidence for how depictive images are represented in the visual brain (**Section 1.2.1.1.**) and evidence for the role of extrastriate areas and frontal regions in constructing depictive mental images (**Section 1.2.1.2.**). Following this, I will consider evidence for visual representations in VWM and examine how individual differences in MI contribute to the visual precision of representations in VWM.

#### 1.2.1.1. The role of early visual areas in depicting mental images

Vital to the development of a depictive theory of MI is the fact that the early visual cortex receives visual information from the retina and forms stable configurations of low-level features (such as, shape, colour, orientation) (Felleman & Van Essen, 1991). The early visual areas, or the early visual cortex, are named as such because V1 is the first to receive visual information in visual perception, and this information is fed up the hierarchy via the ventral and dorsal visual streams. Importantly, V1 is topographically organised, i.e., actual distance in space is mapped to corresponding coordinates on the cortex (Felleman & Van Essen, 1991; Slotnick et al., 2005). This led to the theorisation that if there is activation in V1 during MI, then there is evidence for depictive representations in MI (Kosslyn et al., 2003, 2006; Kosslyn & Thompson, 2003).

The ensuing research using early neuroimaging techniques, including positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), began to provide evidence to support this theory. Both PET and fMRI are adopted in the study of visual systems on the assumption that tasks evoking information processing in the brain corresponds to a change in activation of the specific regions recruited and therefore shifts in blood flow are observed. PET measures change in blood flow in the brain via radioactive tracers and has been argued to be effective in tracking functional specialisation of visual processing (Corbetta, 1993). Similarly, blood oxygenation level-dependent fMRI also tracks haemodynamic responses but via changes in blood oxygenation levels (Kinahan & Noll, 1999). Seminal studies using PET demonstrated that early visual areas activated during perception were also activated when participants were instructed to imagine the stimulus (Kosslyn & Thompson, 2003). In the first study of its kind, lowercase letter cues were presented either below a blank grid (imagery condition) or below a grid with the uppercase letter superimposed (perception condition). In both conditions, an X was placed somewhere on the grid and the participant had to determine whether the X fell on top of the uppercase letter by either viewing the letter and X on the grid (perception condition) or imagining the capital letter and making the judgement (imagery condition) (see Figure 1.1). Importantly, V1 was activated during the imagery condition as well as during perception condition (Kosslyn et al., 1993).

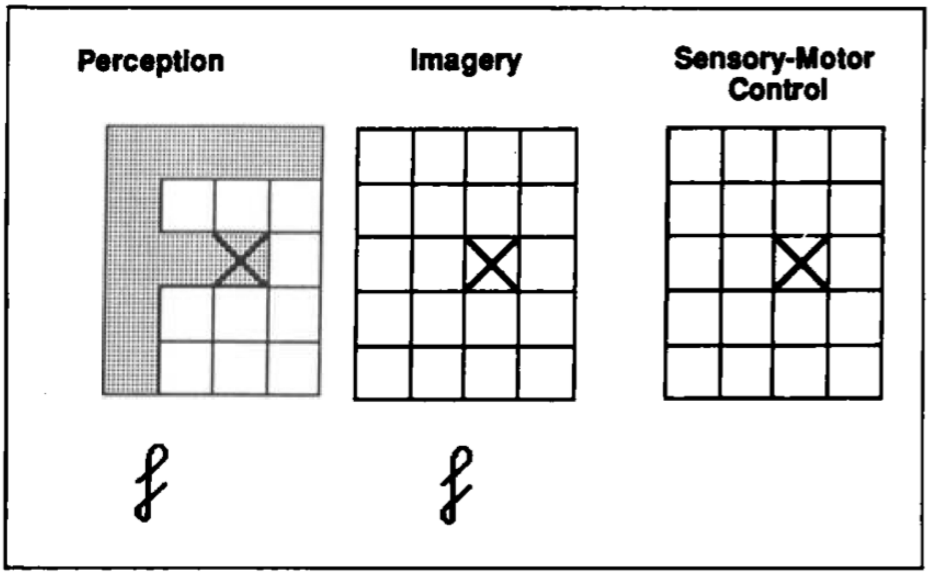



*Figure 1.1:* Perception condition (left), imagery condition (centre), sensory-motor control condition (right). Reproduced with permission from Kosslyn et al. (1993)

In a follow up experiment, participants listened to names of letters and were asked to make a judgement, e.g., does the letter have curved lines? In half the trials, they had to imagine the letter as small as possible and in the other half of the trials they had to imagine the letter as large as they could. Here, it was found that visualising smaller letters activated more posterior regions of the early visual areas. On the other hand, greater activation was found in the anterior region of the early visual areas in trials when larger letters were visualised (see Figure 1.2). Thus, it was argued that more anterior activation corresponding to imagining larger letters suggests that the larger image reflects extension of activation to early visual areas that represent peripheral edges of the visual field. Overall, this was the first study to demonstrate evidence for the hypothesis that generating mental images involves topographically organised areas of the early visual cortex.

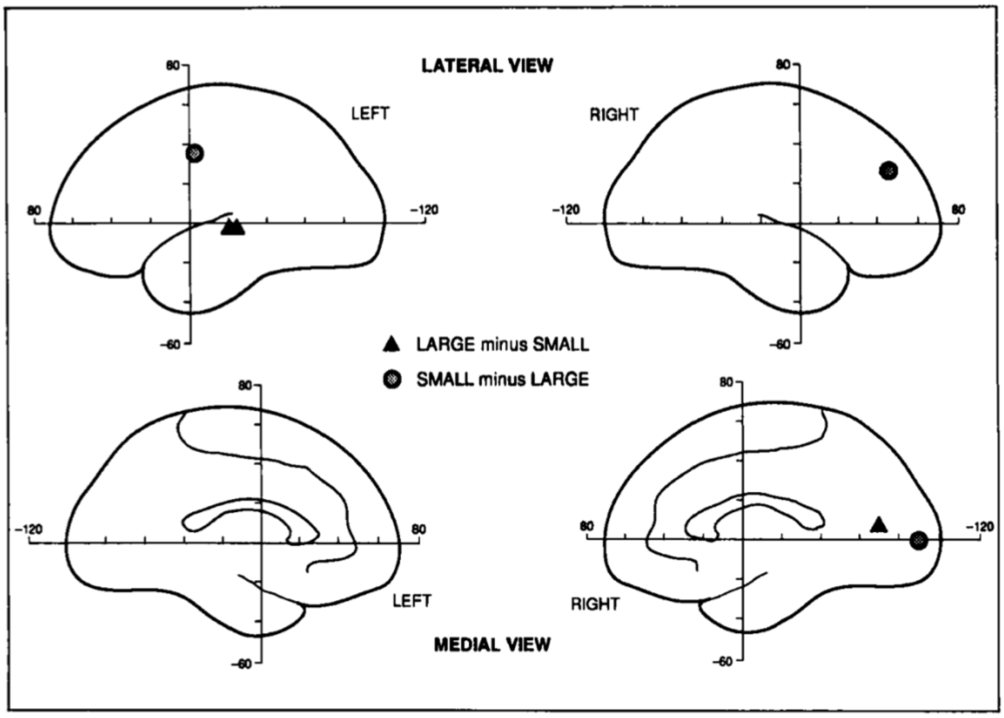



*Figure 1.2:* Bottom-right image represents medial view of right hemisphere activation differences in early visual areas between small and large stimuli trials. Reproduced with permission from Kosslyn et al. (1993)

Subsequent studies conducted over the next decade found evidence for increased V1 activation during MI using PET (Kosslyn et al., 1995, 1996; Shin et al., 1999; Thompson et al., 2001) and using fMRI (Ishai et al., 2000; Klein et al., 2000; Lambert et al., 2002; Le Bihan et al., 1993; Sabbah et al., 1995). However, some studies failed to replicate this finding (Mellet et al., 1995, 2000; Roland & Gulyás, 1995). It could be argued that failure to detect increased activation in the early visual cortex during MI is due to varying task demands. For example, in some studies the control condition was rest (Mellet et al., 1995) rather than visual perception or a simple sensory motor task, therefore imagery during rest in the form of daydreaming cannot be ruled out. Other studies involve complex stimuli, e.g., complex patterns of small dots displayed for only 100ms (Roland & Gulyás, 1995), which arguably could be more difficult to visualise compared to more simple stimuli (e.g., capital letters). Furthermore, some studies found most, but not all participants, demonstrated activation of the early visual areas during MI (Chen et al., 1998; Handy et al., 2004). For example, in Handy et al.'s (2003) study using fMRI, they found 9 out of 14 participants showed V1 activation when they were required to imagine cued pictures that they previously viewed in a perception condition, and 10 out of 14 participants showed V1 activation when freely imagining objects that were cued by listening to the spoken noun. These data arguably demonstrate the importance of recognising individual differences in the extent to which images recruited in MI are visually depictive.

Later research employing more advanced neuroimaging techniques has provided further evidence for shared visual representations in MI and visual perception, however this is not confined to the early visual areas. Studies using fMRI have adopted multivariate pattern analyses (MVPA) whereby classifiers are trained on activation in the early visual areas during visual perception trials and during MI trials. The application of MVPA has been found to decode object-specific information in the early visual cortex. For example, a classifier algorithm could be trained on data from an fMRI session whereby the participant views objects pertaining to specific categories. The classifier therefore learns to map the pattern of activity to an object label. It is then possible to test how accurately the trained classifier can decode activation in the same subset of voxels during viewing of the same object by the same participant in a second fMRI session (Cox & Savoy, 2003; Haxby et al., 2001). Based on this logic, this technique can be applied to examine whether there are shared visual representations between visual perception and MI in the early visual cortex. If decoding accuracy of classifiers trained on MI trials is significantly above chance when applied to activation during visual perception trials, then there is evidence for shared representations.

Studies adopting this technique have demonstrated classifiers trained on activation in the early visual areas during MI trials can successfully decode activation in visual perception trials and vice versa, thus suggesting visual representations are evoked in both visual perception and MI (Albers et al., 2013; Cichy et al., 2012; Koenig-Robert & Pearson, 2019; Lee et al., 2012; Naselaris et al., 2015; Thirion et al., 2006). For example, models trained on voxel-specific spatial frequency and retinotopic location in visual perception, i.e., specific sensory properties, were also found to successfully predict activity in the early visual areas during MI trials (Naselaris et al., 2015). It is also important to note that some have found shared representations using MVPA in late visual areas, e.g., the ventral-temporal cortex, as opposed to early visual areas (Cichy et al., 2012; Reddy et al., 2010), which has led to the argument that while early visual areas can be activated during image generation, it is not a necessary requirement (Reddy et al., 2010). Moreover, while classifiers in these studies can decode activation in trials to a level that is significantly above chance, decoding accuracy is often relatively low when decoding perception activation in MI trials (around 60% accurate compared to around 90% accuracy when decoding perception activation in perception trials) (Koenig-Robert & Pearson, 2019). Therefore, it is important to keep in mind that activity in early visual areas in visual perception does not entirely map onto activity in such areas in MI. This is in line with the suggestion that while visual perception is dependent on bottom-up sensory input, MI is generated from top-down signals (Kosslyn, 2005; Lee et al., 2012).

The finding that the early visual areas are not consistently recruited in MI might be best explained by individual differences in the sensory experience of MI. The sensory experience of MI is most commonly measured using the Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973; Marks, 1995), which involves participants visualising the content of a series of statements and rating how vivid their mental image is from "perfectly clear and as vivid as normal vision" to "no image at all, you only 'know' you're thinking of an object". Studies have initially demonstrated a positive relationship between activation of the early visual areas in MI and VVIQ scores (Cui et al., 2007; Lee et al., 2012), thus suggesting that more vivid MI is associated with greater early visual area activation. Research has since extended such findings by asking participants to rate within-task vividness as opposed to comparing to a general measure of vividness such as the VVIQ. Here it has been demonstrated that higher self-reported trial-by-trial vividness in MI trials is associated with increased activation in the early visual areas (Dijkstra et al., 2017b). Taken together, this research suggests that individual variation in self-reported experience of MI is associated with the neural correlates of visual representations and could account for individual differences in early visual area activation.

To summarise, following the theorisation that MI can evoke depictive representations that are distinct from symbolic representations of general thought (Kosslyn, 1980), preliminary studies showed selective activation of the early visual cortex during MI and visual perception (see Kosslyn & Thompson, 2003 for review). This was supported by more recent evidence employing more advanced neuroimaging techniques and MPVA whereby it was demonstrated that it is possible to decode activation in the early visual areas in MI trials with classifiers trained on visual perception trials and vice versa (Albers et al., 2013; Naselaris et al., 2015). Therefore, this research has provided strong evidence for shared neural substrates in the early visual areas between perception and MI. However, it is important to note that some studies found shared representations between MI and visual perception in late visual areas but not early visual areas (Reddy et al., 2010), which suggests it would be too crude a definition to reduce MI to the activation of a visual representation as in visual perception. Instead, this highlights the role of top-down signals in MI (Lee et al., 2012). Furthermore, some studies have also found that not all participants within one sample demonstrated selective activation of early visual areas during MI, which in turn highlights the importance of considering individual variation in visual representations. In line with this, evidence has shown that individual differences in the subjective, sensory experience of MI is associated with the recruitment of early visual areas in MI (Dijkstra et al., 2017a). Taken together, the evidence outlined demonstrates that while there is no denying that visually depictive representations can be generated in MI, the role of top-down processes cannot be neglected in understanding the neural mechanisms of MI.

#### 1.2.1.2. The role of late visual areas, mid-level, and frontal regions in constructing mental images

As alluded to in the previous section, activation associated with MI is not limited to the early visual cortex, but also extends to extrastriate cortices, or late visual areas, and mid-level regions. To define these areas, the late visual areas include the lateral occipital cortex (LOC) and V4, and mid-level areas extend to the medial-temporal regions (Haxby et al., 2001). The LOC forms the posterior part of the fusiform gyrus and is found to be specifically important in object and shape recognition (Grill-Spector et al., 2001; Malach et al., 1995). Studies have thus compared the role of LOC activation in visual perception and MI. For example, Stokes et al. (2009) trained a classifier on activation recorded using fMRI when participants were imagining either the letter X or O following verbal cues or viewing the letter X or O on the screen. The results showed an overlap between MI and perception in the LOC suggesting that top-down activation is required in the construction of mental images (Stokes et al., 2009), which may be similar to role of the LOC in object recognition during visual perception (Ungerleider & Bell, 2011). Extending these findings, a subsequent fMRI study set out to compare the representations of real-world objects (e.g., a necklace, chair, pen) in MI and visual perception. Using MVPA, this study found greater decoding of perceived objects in early visual areas alongside greater decoding of imagined objects in the extrastriate cortices, including the LOC. This led to the argument that while there are similar neural correlates between MI and visual perception, limited bottom-up input in MI may result in some differential functional activation between MI and visual perception (Lee et al., 2012). Evidence for decodable contents of MI in the extrastriate cortices has since been extended to more detailed and true-to-life stimuli, such as, landmarks (Boccia et al., 2019) and naturalistic scenes (Johnson & Johnson, 2014). With regards to the mid-level neural regions, initial studies found increased activation in the ventral-temporal areas in MI (Ishai et al., 2000; O'Craven & Kanwisher, 2000). Moreover, more recent research employing MVPA to decode the content of imagery has demonstrated that it is possible to decode representations of imagined objects in mid-level regions (Boccia et al., 2019; Breedlove et al., 2020; Cichy et al., 2012; Dijkstra et al., 2017a; Lee et al., 2012; Reddy et al., 2010). Overall, while these findings do not dispute the role of early visual areas in MI, they highlight the importance of late visual areas in MI, as well as differential functional roles of mid-level regions between MI and perception.

Alongside the identification of a role of late visual areas in MI is the emerging argument of the importance of frontal regions. The fronto-parietal network spans the lateral prefrontal cortex and posterior parietal regions and has been shown to underpin goal-directed and executive-control behaviours (Zanto & Gazzaley, 2013). Preliminary neuroimaging studies highlighted selective activation of parietal and frontal cortices during MI (Ganis et al., 2004; Ishai et al., 2000; Mazard et al., 2005; Mechelli et al., 2004). This research has been extended in recent years (Dijkstra et al., 2017a; Winlove et al., 2018) and the recruitment of parietal and frontal cortices has also been found to positively overlap with the trial-by-trial vividness ratings (Dijkstra et al., 2017b). This was argued to represent the importance of parietal-frontal regions in the construction of visual information during imagery. Moreover, findings have demonstrated that it is not only possible to decode real-world objects in early and late visual areas, but it is in fact possible to decode such objects in prefrontal regions during MI

(Ragni et al., 2020). These data support the notion of a top-down mechanism of MI in the absence of bottom-up sensory information.

Some have gone as far as to argue against a purely visual model of MI, criticising the correlational nature of associations between MI and visual areas and citing contradictory evidence in neurological patients (Spagna et al., 2021). For example, it has been shown that individuals with lesions to the occipital cortex have intact MI (Behrmann et al., 1992; Chatterjee & Southwood, 1995), and there is evidence for intact MI in spite of cortical blindness and bilateral damage to the early visual cortex (Zago et al., 2010). In a recent meta-analysis, Spagna et al. (2021) compared activation contrasts in studies whereby MI activation was greater than control conditions (27 experiments) and MI activation was greater than perception conditions (4 experiments) in the regions of interest set out by each study. They found no evidence for increased selective activation in the early visual cortex in MI compared to control conditions, alongside evidence for activation in fronto-parietal networks. In turn, the authors propose an alternative model of MI focusing on a region of the left fusiform gyrus, which they term the Fusiform Imagery Node (FIN) (see Figure 1.3.). It was argued that the FIN might act as a hub for the recruitment of visual representations via the surrounding extrastriate cortices, as well as top-down activation from the fronto-parietal network in constructing the representation. In this sense, individual differences might be dependent on the relative integration by the FIN of surrounding activity (Spagna et al., 2021). While this might appear somewhat alarming considering the extensive evidence outlined thus far, it is important to note that the analysis is based on very specific contrasts. The lack of a difference between early visual cortex activation between MI and control conditions does not refute previous evidence for activation of those areas during MI. Moreover, due to this selection criteria, there were not enough MI vs. visual perception experiments in their study to warrant robust conclusions (only 4 experiments). Nevertheless, the findings further highlight the importance of considering the role of the fronto-parietal network in the generation of mental images.

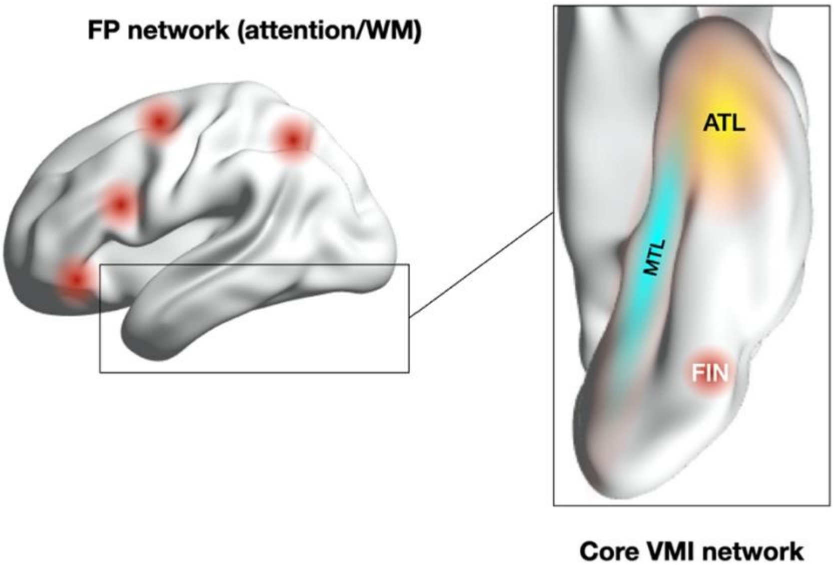



*Figure 1.3:* Depiction of the Fusiform Imagery Node (FIN). Reproduced with permission from Spagna et al. (2021)

If the frontal regions are involved in MI, what is their specific role? When the early visual cortex is engaged in visual perception, activation is projected forward towards the frontal regions via the ventral visual stream (Salin & Bullier, 1995) thereby tracking the processing of low-level (e.g., spatial location, orientation), mid-level (object recognition) and high-level features (semantics) of what we see in front of us, and activity is then projected back again to the early visual cortex to refine and stabilise representations (Albers et al., 2013). Given that MI is characterised as a mental representation in the absence of sensory input, it has been conceptualised that the re-activation of information stored in memory to generate a mental image comprise top-down activation of high-order areas, which in turn projects back to the early visual cortex to form a visually depictive representation (Kosslyn, 2005).

Recent research adopting causal modelling techniques has provided evidence for the notion that MI is modulated by a top-down flow of activation from the frontal areas to posterior early visual areas; thus, the reverse of the direction of connectivity observed in visual perception (Breedlove et al., 2020; Dentico et al., 2014; Dijkstra et al., 2020). For example, a recent study recording electrophysiological data with ms-by-ms precision found a negative relationship between perception activation models trained on early time points and those same time points in the MI trials, which indicates that frontal regions were activated earlier than visual areas in MI (Dijkstra et al., 2020). The notion of reversed directionality in MI compared to perception has also been supported by research that has demonstrated it is possible to fit a generative feedback network model to neuroimaging data during MI (Breedlove et al., 2020). Taken together, the evidence supports the notion that activation of the fronto-parietal network is reversed in MI compared to perception and thus suggests a pivotal role of the frontal regions in the generation of visual representations.

To summarise, the research implies different functional roles of the mid-level visual regions in MI compared to perception, which suggests top-down modulation of MI. In support of a top-down mechanism in MI, there is evidence for selective activation of parietal and frontal regions during MI and decoding representations in prefrontal areas of imagined objects, landmarks, and naturalistic scenes. Consequently, a recent meta-analysis concluded that the neural model of MI should be revised to emphasise the role of the fronto-parietal networks (Spagna et al., 2021). Thus, the evidence shows that activation in brain regions underpinning high-order cognitive functions is prominent in MI and therefore aligns with the complex and flexible way in which we use MI in everyday thinking and learning. This perspective is directly applied in **Chapter 4** by examining the neural correlates of the role of MI in VWM in posterior electrodes, i.e., over the visual areas, and frontal electrodes.

#### 1.2.1.3. MI as an explanation for individual differences in VWM

Central to our understanding of the format of MI is the investigation of the role of MI in VWM, however few studies have examined this relationship and the divergent literatures of MI and VWM rarely communicate. If we are to understand how MI presents in groups with working memory impairments, specifically, children with ADHD, then we must fully establish how individual differences in MI impact VWM. The function of VWM seems to echo that of MI in that visual information can be stored temporarily, focussed on at will and reactivated as needed (Baddeley, 2003; Baddeley & Andrade, 2000; Cowan, 2001; Logie, 1995). The neural correlates of VWM, with respect to electrophysiological signatures and functional, selective activation, have been investigated extensively to determine the mechanisms underlying capacity limits in VWM. Namely, it is well-known that individuals can retain and manipulate up to 3-4 items in VWM, despite there being extensive individual differences in that this range spans from 1 to 7 items (Vogel & Machizawa, 2004). The extent to which individual differences in VWM capacity are dependent on the visual precision of representations has been investigated in three parallel strands of research. The first has examined the role of frontal regions compared to early visual areas in VWM maintenance (Miller & D'Esposito, 2005; Sreenivasan et al., 2014). Another branch of research has assessed the neural correlates of the precision and capacity at which visual information is maintained in VWM (Adam et al., 2018; Luria et al., 2016; Machizawa et al., 2012, 2020; McCollough et al., 2007; Vogel & Machizawa, 2004). Finally, some behavioural studies have investigated how individual differences in MI are associated with VWM ability (Keogh & Pearson, 2011, 2014). The following section will draw together the findings presented in these three strands of research and present the argument that a VWM paradigm can provide a novel opportunity to fully examine how individual differences in MI support the precision and capacity of VWM.

Research over the past two decades has sought to clarify the functional activation associated with VWM using fMRI; however, the evidence is not clear cut. The dominant finding in this literature suggests the involvement of the lateral prefrontal cortex (lPFC), however the specific functional role is up for debate (Sreenivasan et al., 2014). It was initially thought that selective activation of the lPFC must be attributed to the encoding of sensory information during VWM maintenance within a delay period (Constantinidis et al., 2001), however it is now more commonly accepted that the top-down activation of the lPFC modulates the extrastriate cortex in maintaining visual representations (Lorenc et al., 2015; Miller & D'Esposito, 2005; Sreenivasan et al., 2014). Moreover, the importance of the visual areas in VWM maintenance is supported by MVPA analyses of fMRI data. That is, it has been demonstrated that stimulus content held in mind during the delay period of a VWM task can be decoded in the early visual areas (Albers et al., 2013; Harrison & Tong, 2009). Thus, there appears to be crossovers in the contribution of frontal versus visual regions in MI and VWM. This is supported by recent findings showing that classifiers trained on activation in the early visual areas during MI trials successfully decoded stimulus representations in VWM trials and vice versa (Albers et al., 2013). Moreover, it was also shown in Albers et al.'s (2013) study that although selective activation of parietal and prefrontal regions was noted in both processes, decoding accuracy for stimulus content was at chance in these areas. Therefore, this research supports the notion of the role of frontal regions in mediating content-specific representations in the early visual cortices in both MI and VWM.

The suggestion that content-specific information is represented in the visual areas is supported by a neighbouring line of enquiry using electroencephalogram (EEG). In a seminal paper, Vogel and Machizawa (2004) found neural correlates of VWM maintenance were characterised by an event related potential in the occipital and parietal electrodes, termed contralateral delay activity (CDA). CDA is found to be modulated as a function of the number of items held in mind up to 4 items. The finding that CDA can index VWM capacity has since been replicated (see Luria et al., 2016 for review), and there is evidence for individual differences in that greater CDA amplitude is denoted in individuals with good VWM compared to those with poorer VWM (Adam et al., 2018). Furthermore, this research has been extended to demonstrate that it is possible to examine the precision at which representations are held in VWM via CDA amplitudes. Researchers applied an orientation-discrimination paradigm to not only discriminate between CDA amplitudes associated with increasing set size, but also those associated with coarse (45°) and fine (15°) orientation discriminations. Here it was found that at smaller set sizes, there was greater CDA amplitude in fine orientation discriminations compared to coarse. Thus, it was interpreted that at lower capacities, individuals are able to maintain representations of fine precision and that the CDA amplitude can reflect both precision and capacity of maintained representations (Machizawa et al., 2012). What is important to note here is a distinction in terminology; what might be described as precision in VWM literature could be described as the visual quality of a mental image through the lens of MI research. However, what is still unclear, and prevents the opportunity to make such a conclusion, is how the visual quality of such representations can impact VWM performance and modulate CDA amplitudes.

While the role of MI in VWM has not been directly examined, performance on MI tasks has been compared to performance on VWM tasks. Using a binocular rivalry paradigm, researchers have derived a MI sensory strength measure (Pearson et al., 2008) and have subsequently examined whether strategies recruited in VWM differ between individuals demonstrating different levels of MI sensory strength (Keogh & Pearson, 2011, 2014). Firstly, it has been demonstrated that MI strength is positively associated with VWM capacity, but not iconic (short-term memory of one item) or numeric memory (remembering number strings) (Keogh & Pearson, 2011). Moreover, modulating background luminance, which has been previously shown to disrupt visual representations held in mind (Baddeley & Andrade, 2000; Pearson et al., 2008), was found to impact VWM performance only in those that scored highly on the MI strength measure. This was interpreted as indicating that only “good imagers” adopted an MI strategy in VWM (Keogh & Pearson, 2011, 2014).

This preliminary evidence supports the notion that the recruitment of visual representations in VWM is dependent on individual differences in the sensory strength of imagery. However, associating VWM performance with the MI strength measure may not entirely characterise the visual quality of representations in VWM. Firstly, because the nature of this investigation is based on associating performance on two separate tasks, we cannot draw conclusions regarding visual quality of representations recruited *within* the VWM task and thus we cannot fully determine how MI might support VWM abilities. Secondly, recent findings have suggested that while the imagery priming effect using the binocular rivalry paradigm is found at the group level, this is not found reliably at the individual level (Dijkstra et al., 2019). Taken together, to fully examine how individual differences in MI contribute to VWM ability, further research is required to assess how the visual quality of representations impacts VWM performance and neural correlates of VWM.

### 1.2.2. Summary and open questions

Principal theorisation of MI emphasised two crucial attributes. Firstly, that MI is not a unitary construct or singular function but is in fact a multi-faceted construct (reviewed in the next section). Secondly, that visually depictive representations can be generated in MI (reviewed above). Based on what is known about the visual system when we visually perceive a stimulus, research to-date has compared the nature of representations in visual perception to those generated in the absence of visual input. Collectively, this provided clear evidence that individuals *can* generate visually depictive images, and this is subject to individual differences in the sensory experience of MI. More recent research has revisited the role of the extrastriate and frontal cortices to examine how mental images are constructed; this has in turn highlighted that top-down activation of frontal cortices is vital in MI. In light of this, it is now apparent that the quest to equate representations in visual perception and MI has reduced MI down to the presence or absence of a visual representation, which neglects the complexity of the sensory experience and function of MI. It is therefore more appropriate to adopt a theoretical model that considers the role of the fronto-parietal network and is also more aligned to how we generate, maintain, manipulate, and inspect mental images to support our everyday thinking. By reapproaching the study of MI with the view that MI is a complex, multi-faceted construct, we can address the gaps in our understanding of the mechanisms of MI. The following section will review evidence of how a separable-component model of MI can provide a framework to address these gaps in knowledge and this is fully investigated in **Chapter 2.**

In parallel, research has examined functional selectivity and the electrophysiological correlates that underpin the maintenance of visual representations in VWM, and this has shown that representations can include a varying number of items represented at varying levels of precision. While drawing parallels between MI and visual perception has furthered our understanding of the format of representations in MI, the findings outlined in **Section 1.2.1.3**. suggest that shared mechanisms are apparent between MI and the high-order cognitive function of VWM. This might lead to the assumption that visual mental images are recruited in both MI and VWM and that MI plays a role in supporting VWM, however, preliminary findings suggest only those with strong MI recruit this strategy in VWM. If this were true, if greater VWM ability is underpinned by the visual quality of mental images, then this would have important implications for the way in which we conceptualise MI and would further elucidate the role of MI in memory and goal-directed behaviour. In turn, it might be anticipated that poorer MI plays a role in the VWM impairments observed in children with ADHD. This would present a unique opportunity to intervene with VWM difficulties in this group. However, MI abilities have not been characterised in children with ADHD, and the relationship between MI and VWM in both the ADHD group and typically developing children is unknown. The relationship between MI and VWM is fully investigated in primary aged children (age 6-11 years) in **Chapter 2** and children with ADHD in **Chapter 3**.

Furthermore, research thus far has compared ability and neural correlates on MI vs VWM tasks, which has led to assumptions about the role of MI within VWM. However, the investigation of how individual differences in MI within a VWM task has not been directly tested. In **Chapter 4**, the role of MI in VWM is fully investigated by examining how individual differences in MI impact the precision and capacity at which visual information is held in mind within VWM via neural correlates of VWM maintenance, as well as examining the role of the frontal cortices in generating visual mental images to support VWM.

### 1.2.3. MI as a multi-faceted construct

While one strand of research in MI has focused on determining the format of representations, the other has investigated the cognitive function of MI. Following the early theorisation of the sub-component model of MI, few studies have applied this model in MI research. As the sensory experience of MI is variable, research has thus far largely focused on using neuroimaging techniques to uncover the format of representations recruited in MI, which has revealed evidence for visually depictive mental images as outlined in the section above. However, this neglects the wealth of information that can be provided by investigating psychological phenomena using carefully crafted behavioural tasks. While the binocular rivalry paradigm has been adopted in MI research as a behavioural measure of the sensory strength of MI (Pearson et al., 2008; Pearson & Kosslyn, 2015) this does not account for the complexity of MI abilities. A key observation in the early development of a theoretical model of MI was that MI is not a unitary construct: rather, it comprises multiple sub-abilities that are distinct from one another and influenced by extensive individual variation (Kosslyn, 1980; Kosslyn et al., 2006). Such a definition is more akin to the way in which we experience MI in everyday life, i.e., a highly visual image is not always the most efficient. For example, a highly visual image might be useful when trying to visualise a descriptive scene whilst reading a story. On the other hand, envisioning how your luggage might fit in the boot of a car is likely more dependent on your ability to transform spatial relations in mind. Some abilities might be dependent on both strong visual and spatial features, such as describing the best route when giving directions. Nevertheless, research adopting a sub-component model to investigate the mechanisms of MI is limited.

In addition, the investigation of how MI develops throughout childhood in the context of a separable-component model and the investigation into the format of representations in children's MI have been almost entirely neglected. From a development perspective, there is great value in adopting a multi-component framework of cognition as opposed to adopting general and unitary perspectives (Karmiloff-Smith, 1994). Given how little is known about the development of MI abilities in childhood, the application of a subcomponent-model presents a clear framework to address this gap. Characterising the development of MI is not only imperative to our understanding of the emergence of this ability in both typical and atypical development, but also has implications for the contribution of MI to other cognitive abilities vital for learning and problem-solving.

#### 1.2.3.1. Evidence for a sub-component model of MI

The sub-component model of MI is made up of key sub-abilities: primarily, the ability to generate a visual image (image generation), the ability to hold a visual image in mind (image maintenance), the ability to transform an image (image transformation) and the ability to shift attention across or inspect an image (image scanning) (Kosslyn, 1980; Kosslyn et al., 2006; Kosslyn et al., 1984). It is important to note here that the sub-component model of MI is not dissociated from a depictive theory of MI; they are integrated in that sub-component measures pinpoint individual differences in the generation, maintenance, transformation, and inspection of visually depictive mental images. While other sub-abilities have been described, the sub-component model focusing on the four components has provided a useful framework for characterising MI abilities (Dror & Kosslyn, 1994; Kosslyn et al., 1990; Wimmer et al., 2015, 2017). In the following section, I will outline evidence for each of the sub-components of MI and discuss how this framework can be applied to capture individual differences in MI abilities from childhood through to adulthood.

In a seminal paper, Kosslyn and colleagues were the first to examine the underlying structure of MI by assessing performance on measures of multiple sub-abilities, including those that formed the sub-component model. A lack of association between each of the measures was found in the participant data and in a computational model simulating the parameters of MI, which provided evidence for the theory that the structure of MI is made up of distinct sub-abilities (Kosslyn et al., 1984). This finding was subsequently supported by neuroimaging research. Specifically, using PET it was found that different brain regions predicted performance in different MI components, which was concluded to suggest components of MI are distinct from one another (Kosslyn et al., 2004). Alongside evidence for dissociated activation, there was also evidence for some overlap between regions activated; primarily in the occipital-parietal sulcus, medial frontal cortex, and early visual areas in the occipital cortex. Considered alongside the evidence presented in **Section 1.2.1.**, this reflects what is now known about the functional roles of both visual and frontal regions of the brain in MI. In the remainder of this section, I will review evidence for the each of the components, how they develop throughout childhood and how they relate to one another throughout development while outlining the current gaps in knowledge.

##### *1.2.3.1.1. Image generation*

At the core of the sub-component model of MI is image generation; the ability to generate a mental image in the absence of sensory input, which is argued to involve reactivating visual information from long-term memory (Kosslyn et al., 2006). Image generation is arguably the most researched component as most behavioural and neuroimaging studies of MI require an individual to generate an image and this is often defined as the MI task. Notable examples of this are outlined in the studies described in **Section 1.2.1**; paradigms used in neuroimaging research examine activation associated with instructing an individual to generate an image. This raises two issues. Firstly, it means that there are many different paradigms used to measure the ability to generate an image, which makes it difficult to compare across studies. Secondly, as most studies only include the generation of an image as their assessment of MI ability, there is little investigation of how this component is related to the other components of MI.

Studies can largely be differentiated by paradigms that include a delay period whereby participants are simply instructed to imagine a stimulus (e.g., Dijkstra et al., 2017; Lee et al., 2012; Stokes et al., 2009) and paradigms that tap into the accuracy in which individuals generated an image (e.g., Kosslyn et al., 1990, 1993; Wimmer et al., 2015). Preliminary tasks involved questioning participants on known artefacts. For example, the animal's tails test involves asking participants to judge whether an animal's tail is long in proportion to its body or whether it has floppy or upright ears (Kosslyn et al., 1977; Pearson et al., 2013; Policardi et al., 1996; Zeman et al., 2010). In the same vein, the commonly applied mental clocks task requires participants to determine whether the angle between the hour and minute hand is smaller than 90° (Berryhill et al., 2007; Paivio, 1978; Rosenbaum et al., 2004; Trojano, 2000). The principle of these tasks is that it is assumed that participants need to generate a mental image to answer correctly. However, it could be argued that there are various strategies involved, for example, individuals may have abstract knowledge of an animal feature which they can refer to. Other tasks appear to be more successful in isolating the ability to generate a mental image. For example, the capital letters task, described in detail in **Section 1.2.1.1.** and outlined in Figure 1.1., requires participants to generate an image of a capital letter on a presented grid outline based on a lowercase letter cue and detect whether an X would fall on the imagined letter. While imagining a capital letter might not require a mental image, having to imagine the letter within the specific shape of the grid arguably does require the generation of a visual image to answer correctly. Other studies have compared participants ability to imagine non-object or abstract shapes to known objects and found differential activation of the early visual areas between the conditions (Mazard et al., 2005). Thus, this research suggests that while it is possible to isolate the ability to generate a visual image, careful consideration of task paradigms is required.

#### *1.2.3.1.2. Development of image generation*

To date, there are just three studies that have investigated how the MI components of the sub-component model develop throughout childhood (Kosslyn et al., 1990; Wimmer et al., 2015, 2017). In the first study of its kind, Kosslyn et al. (1990) examined ability in each component of MI in children aged

5, 8 and 14 years and adults. The image generation task was a modified version of the capital letters task in that participants had to detect whether two Xs fell onto an imagined capital letter and letters were separated into simple letter and complex letter trials. Firstly, findings regarding developmental progression suggested that adult-like ability in image generation was not present until age 14 years. Further analysis focused on the difference between simple letters, i.e., fewer segments such as ‘C’, or complex letters, i.e., more segments such as ‘G’, which the authors argued allowed for comparison between a simple image or more complex image. Image generation trials were also compared to perception trials, whereby participants viewed the capital letter on the grid and determined if two X’s were presented superimposed onto the capital letter or not. Findings suggest longer response times in generating a complex image compared to the perception condition, which was interpreted to demonstrate that additional processes are required in image generation compared to perception. There was also a complexity effect in that in all age groups apart from the 5-year-olds, there were longer response times (RTs) for complex letters compared to simple letters. While it was also expected that error rates would be higher in complex compared to simple letters, this was found only in the 5-year-olds and 8-year-olds but not the 14-year-olds and adults. This raises an important issue of whether the differences in error rates are dependent on differences in image generation or differences in familiarity and knowledge of uppercase and lowercase letters. The use of letters arguably confounded results given that the 5-year-old group were excluded from most analyses. Secondly, this evidence might appear to contradict what we now know about the timing of MI, i.e., images can be generated in a little as 200ms (Dijkstra et al., 2018). Therefore, the longer response times in the complex letter condition might more likely depend on the requirement to inspect a more complex image to determine the answer. Taken together, further development of age-appropriate tasks that negate the requirement of previous knowledge and tasks that isolate image generation are required.

Recent research has developed more age-appropriate tasks suitable for children from age 4 years (Wimmer et al., 2015). In their image generation paradigm, participants had to memorise the location of a known object (e.g., a vase) (see Figure 1.5A). They then completed a 30 second distractor task, whereby 4-year-olds had to count, and older children and adults were required to complete mathematical calculations that were increasingly complex with age. A target object was then presented (e.g., a flower) and participants had to determine whether the target object was presented in the same location as the previously memorised object. New objects were presented either superimposed, near or far from the previously memorised object. In contrast to Kosslyn et al. (1990), it was reported that image generation ability was present from age 4 years with adult-like ability reached around 10 years of age. A second experiment was conducted to assess the precision at which mental images are generated and maintained. In these image generation and image maintenance tasks, the trial sequence was the same except that participants were required to drag and drop the sample object to its original location and then following a distractor task (image generation) or short delay (image maintenance), the participant was required to move the probe object to the location of the sample object. Precision was assessed as the overlap between the location of the placed object and the location of the sample object. Here, it was found that children age 6 years generated mental images with similar precision to adults and children age 8 years maintained mental images with similar precision to adults.

While this task addresses the previous confounding effect of age, the use of known objects combined with varying locations, as opposed to varying visual details of the image, could encourage verbal strategies. For example, verbally remembering that a picture of a vase was left of centre when determining the location of the target object is possible in this task. At this point, it might be argued that this is also a confound in the neuroimaging studies listed above that include known stimuli (e.g., Lee et al., 2012; Ragni et al., 2020). However, neuroimaging studies have much shorter paradigm presentations, e.g., in one trial, a cue could be presented for just 500ms with only a 2 second period to generate the image before moving onto the next trial (Lee et al., 2012). In the Wimmer et al. (2015) task, one trial equates to just under 60 seconds depending on RTs. Such timings are more appropriate for examining the developmental progression of image generation given that developmental differences in speed of processing are less likely to be a confound (Kail, 1991). However, in the context of a longer paradigm, it is necessary to control for the propensity to use non-imagery strategies to fully isolate the ability to generate a visually depictive mental image. Secondly, there is the obvious ecological strength that using known stimuli is more akin to the kind of images we might generate in real-life. However, when in the context of longer trial sequences, it is difficult to determine the extent to which MI strategy is used in this task.

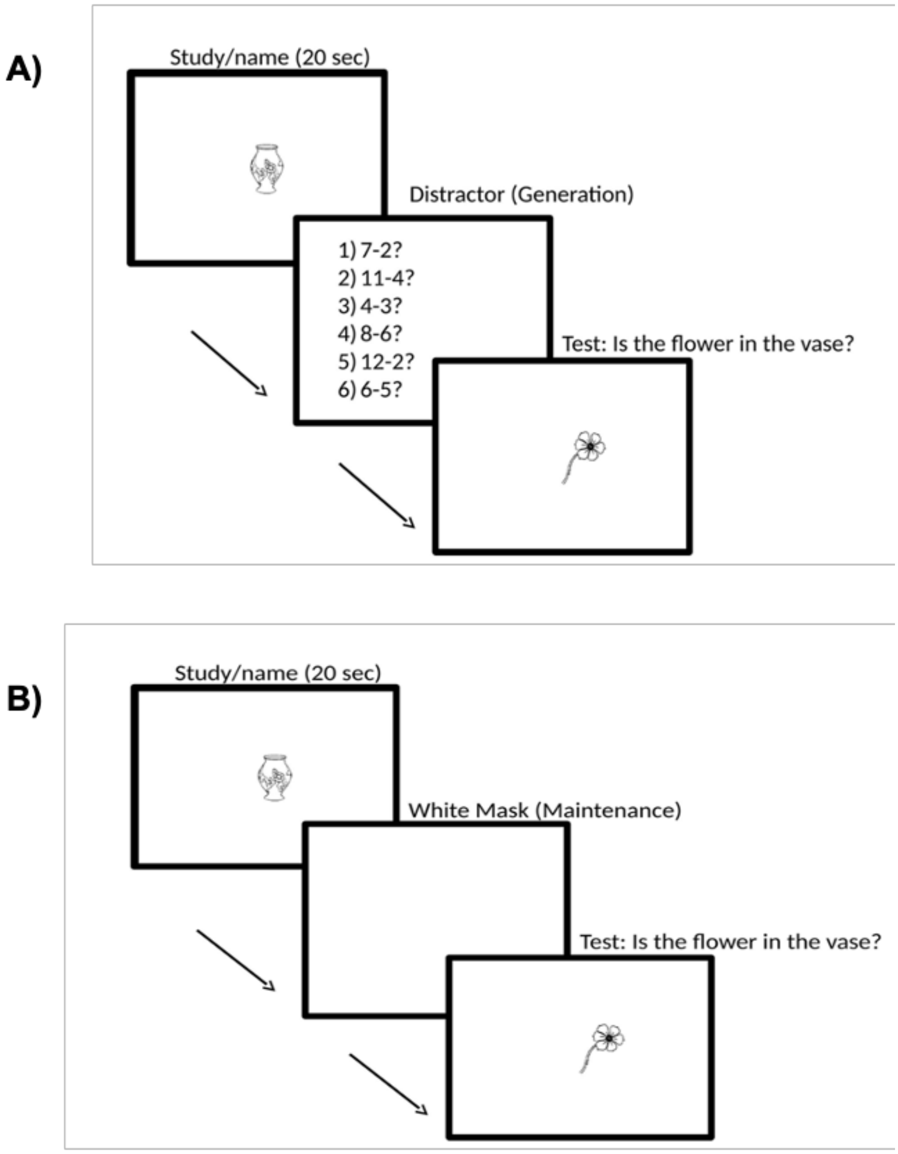



*Figure 1.5:* A) Image generation trial sequence (example of superimposed trial): participants memorise the object and its location (20s), participants 6 years and above complete a math calculation distractor task (4-year-olds complete counting task), target stimulus is presented, and participants have to determine if the target object is in the memorised object. B) Image maintenance trial (example of superimposed trial): exactly the same trial sequence except that a white mask is in place of the distractor task and presented for either 500ms or 3000ms. Reproduced with permission from Wimmer et al. (2015)

It should be noted that there are other studies that have employed MI tasks pertaining to the sub-component model with childhood samples. Again, not all components are assessed and different tasks are implemented, including tasks that require participants to reproduce memorised symbols (Commodari et al., 2020) and tasks that require the reproduction of an object associated with a letter cue (Guarnera et al., 2019). However, these studies only investigate how performance on the tasks predicts performance on measures of learning or school readiness, rather than examining individual differences in task performance or developmental progression of image generation ability. Therefore, it is not possible to gauge how sensitive these tasks are to measuring the ability to generate a mental image. Overall, in light of evidence of confounded image generation tasks (Kosslyn et al., 1990), possible recruitment of verbal strategies (Wimmer et al., 2015) and contrasting findings regarding developmental progression, it is not possible based on the current evidence to make conclusions regarding the development of image generation with respect to generating visually depictive mental images.

*1.2.3.1.3. Image maintenance*

In the adult literature, image maintenance is often intertwined with image generation because many paradigms assessing the ability to generate an image involve holding that image in mind for a number of seconds (e.g., Boccia et al., 2019; Kosslyn et al., 1990; Lee et al., 2012). Correspondingly, in the developmental literature of maintenance in VWM, rather than MI, has been investigated extensively (Cowan, 2014; Simmering, 2012). Findings have suggested the precision at which information is held in mind during VWM is still developing into early adolescence (Burnett Heyes et al., 2012). However, there is little research investigating the ability to maintain an image in the context of the sub-component model of MI. Assessment of VWM maintenance in childhood rarely aims to establish the mechanisms of this ability, and more often aims to determine how VWM maintenance predicts other abilities such as reasoning and measures of academic achievement (Berg, 2008; Best et al., 2011; Bull et al., 2008). Therefore, there is a gap in the literature pertaining to individual differences in the maintenance of visually depictive mental images throughout development. To fully examine the components of MI, how they develop throughout childhood and how they relate to one another it is advantageous to include sensitive measures of both image generation and image maintenance.

#### *1.2.3.1.4. Development of image maintenance*

Kosslyn et al. (1990) included an image maintenance task to track ability to retain a mental image over developmental time. The task used in this study was a variant of the image generation task in that it also required identifying whether two X marks fell in certain positions on a blank grid. However, rather than generating an image of an uppercase letter following a lowercase letter cue as in the image generation task, participants first viewed a pattern of randomly filled squares that filled 20% of the grid and they then had to determine whether the two X's presented on the subsequent blank grid fell onto the memorised pattern. The amount of time maintained was also manipulated in that in half the trials, participants had to maintain the image for 500ms on a smaller grid (4x5 squares) and in the other half they maintained the image for 3000ms on a larger grid (5x7 squares). Findings demonstrated higher error rates and longer RTs in the maintenance of the larger grid for 3000ms. However, the effect on RT was not evident in the 5-year-olds and the effect on error rates was not evident in adults. Moreover, developmental progression was evident in that error rates increased with age (5, 8, 14 years and adults tested), however post-hoc comparisons between age groups were not reported therefore the age at which adult-like abilities are reached is unclear.

Conversely, Wimmer et al. (2015) devised a paradigm that followed the same sequence as their image generation paradigm but with only one difference; instead of the distractor task, participants either maintained the image for 500ms and 3000ms (see Figure 1.5B). With regard to time maintained, Wimmer et al. (2015) found no differences between 500ms and 3000ms in any of the age groups (4, 6, 8, 10 years and adults tested). Progression over development was found in that adult-like ability was reached at around 8 years of age. By differentiating the image generation and image maintenance paradigms on only one parameter, this rules out other potential confounds, such as difficulty of stimuli between the two tasks. Taken together, further clarity is needed on how the ability to maintain a visually depictive mental image develops throughout childhood.

#### *1.2.3.1.5. Image transformation/mental rotation*

Image transformation is defined as the manipulation of an imagined object or anticipating what would be seen if an external force manipulated an object (i.e., mental rotation) (Kosslyn et al., 2006). The ability to transform an image is most often measured by a mental rotation task. Central to our understanding of the ability to manipulate information in mind are the classical studies on mental rotation. Mental rotation tasks require participants to discriminate between two or more stimuli by rotating a target stimulus in mind. The original paradigm from Shepard and Metzler (1971) involves comparing two 3-dimensional (3D) cubed stimuli to determine whether the second, rotated stimulus is mirrored (i.e., different) or not mirrored (i.e., same) compared to the first stimulus. Specifically, evidence for a linear time-degree of rotation effect whereby RTs increase as a function of increasing degree of rotation has been replicated extensively since its original conception (Borst et al., 2011; Ganis & Kievit, 2015; Shepard & Metzler, 1971). Thus, suggesting that individuals rotate an image in mind. This evidence is supported in neuroimaging research. Firstly, fMRI research has indicated differential activation in the early visual cortex for differentially orientated stimuli (Klein et al., 2004). Moreover, evidence from electrophysiological research has also denoted a rotation-related negativity component in the occipito-parietal cortex that maps this effect in that the amplitude increases with increasing degree of rotation (Riečanský et al., 2013; Riečanský & Jagla, 2008). Taken together, this research supports the notion that the increase in RT with increasing degree of rotation reflects the incremental transformation of a depictive representation (Borst et al., 2011; Kosslyn et al., 2006; Shepard & Metzler, 1971). The term image transformation was adopted in early theoretical development of a sub-component model (Kosslyn, 1996; Kosslyn et al., 1984), however mental rotation has been researched extensively in the context of spatial cognition literature and the term image transformation is less common. For completeness, I will use the term mental rotation for the remainder of the thesis.

*1.2.3.1.6. Development of mental rotation*

A precursor to the ability to rotate 2D and 3D objects has been found in infants as young as 3 months old (Hespos & Rochat, 1997; Moore & Johnson, 2011), that is, mental rotation skill is interpreted as preferential looking behaviours between familiar and mirrored objects (Moore & Johnson, 2011). With regard to demonstrating evidence of rotating an image as in adult studies, research has found this ability is present from age 4 (Estes, 1998; Marmor, 1975). However, there are some discrepancies as to whether this ability is present in 4 or 5-year-olds (Frick, Hansen, et al., 2013; Wimmer et al., 2017), which suggests there is still development of this skill during early childhood. Studies examining the development of mental rotation in childhood have adopted the structure of the original paradigm from Shepard and Metzler (1971) and modified to include different types of stimuli which are presented in 2D rather than 3D. Preliminary findings involving a pair of 2D bear stimuli found 4-year-olds demonstrated the linear-time degree of rotation effect (Marmor, 1975), however a later study with this task found chance performance in 4- to 6-year-olds (Dean & Harvey, 1979). Since then, various adaptions of the task have been implemented in order to control for potential task difficulty confounds that could be masking developmental progression (Estes, 1998; Frick, Hansen, et al., 2013; Lütke & Lange-Küttner, 2015; Rosser et al., 1985). For example, research has shown evidence of mental rotation skill from age 4 using a task with 2D human figures instead of 3D cubes (see Figure 1.6) (Estes, 1998), and further investigation has shown children perform better in tasks involving animal stimuli compared to cubes (Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011). Thus, while there may be an indication of mental rotation ability from early infancy, the stimulus-type in tasks adopted to characterise mental rotation in childhood is important to consider in the interpretation of results.

While many studies have assessed mental rotation abilities in childhood, only two have assessed this ability in the context of the sub-component model of MI. Initially, Kosslyn et al. (1990) adapted Shepard and Metzler’s task to include 2D cube forms as opposed to 3D. Stimuli was produced on a 4x5 grid, as used in their image generation and image maintenance tasks. The first stimulus was presented upright (0°), and the second stimulus was presented at 36° increments from 0° to 180°. In half the trials, the second stimulus was a mirror image of the first and in the other half it was identical apart from the rotation. Participants had to determine whether the two stimuli presented were the same. Findings show that while younger children were slower to respond than adults, all age groups demonstrated RT increased as a function of distance rotated, thus supporting earlier findings that mental rotation skill is present from age 5.

On the other hand, Wimmer et al.'s (2017) recent study used Estes (1998) task as pictured in Figure 1.6. In line with common mental rotation task parameters, the second stimulus was either the same (same arm lifted) or different (opposite arm lifted) and it was presented at 30° increments from 0° to 180° compared to the first upright (0°) stimulus. In line with Kosslyn et al.'s (1990) findings, they reported decreasing response times with age. Moreover, an analysis of slopes revealed that children from age 4 demonstrated the linear time-degree of rotation effect, however a main effect of age revealed that 4-year-olds differed significantly from all other age groups, whereas 6-year-olds demonstrated adult-like slopes. Taken together, the findings demonstrate that children are able to rotate an image in mind from age 4, however there appears to significant development in this skill until age 6.

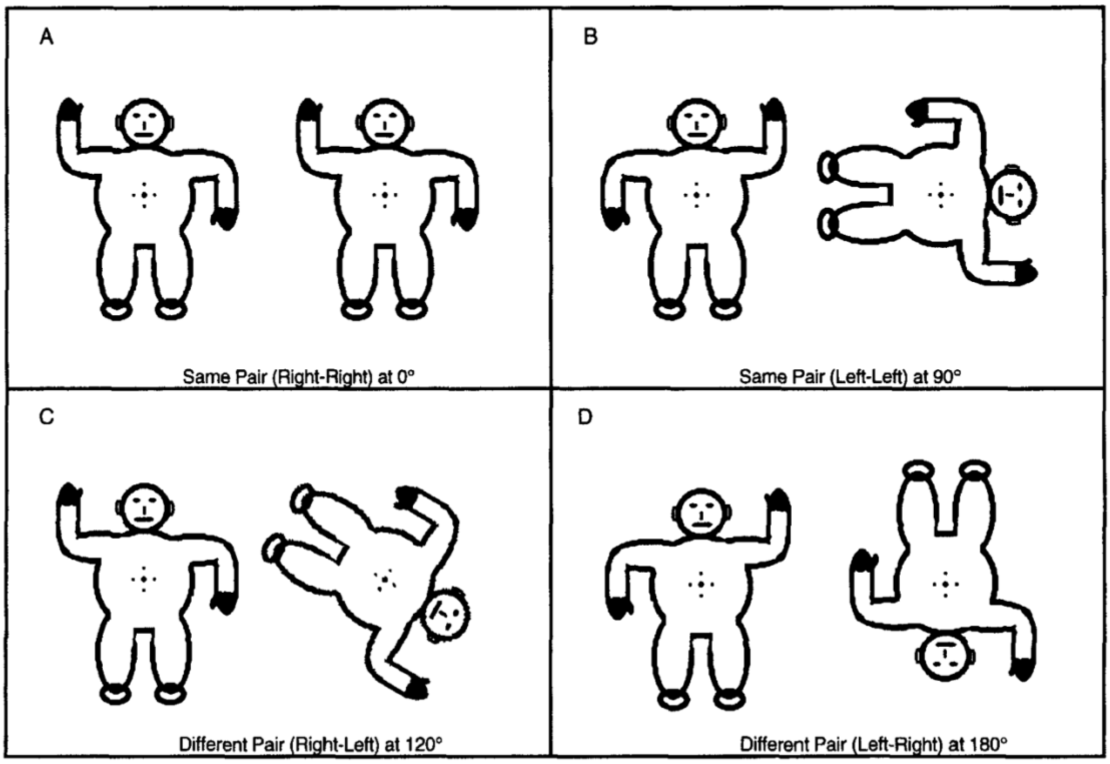


*Figure 1.6:* Sample stimuli of four different stimulus pairs from a 2D mental rotation task suitable for use with children. Reproduced with permission from Estes (1998)

#### *1.2.3.1.7. Image scanning*

Image scanning is defined as the ability to shift attention across a mental image (Kosslyn et al., 2006). If you're stood in a furniture shop trying to picture how a new sofa might fit into your living room, you might have a picture of the room in mind and you might inspect the areas of the room to figure out where in the room it might fit. The ability to inspect different parts of an image is termed image inspection. Importantly, it is argued that during image inspection, spatial features such as distances, orientation and configurations of objects, are encoded (Kosslyn et al., 2006). This phenomenon is often measured using an image scanning task, whereby participants are required to memorise a scene and shift their attention between objects within the scene. Conceptually, image inspection can involve image generation and vice versa in that one can inspect a part of a generated image. While both have been shown to activate frontal regions, there are also dissociations between the brain areas that predict performance on each of the tasks (Kosslyn et al., 2004). Therefore, by deriving sensitive dependent variables, it is possible to use an image scanning task to isolate the ability to shift attention across the spatial properties of a mental image.

Evidence for the ability to shift attention across a visually depictive image is demonstrated in linear relationships between the distance scanned and response times (Borst & Kosslyn, 2008; Borst et al., 2006; Borst & Kosslyn, 2010; Finke & Pinker, 1982; Kosslyn et al., 1978). There are two key paradigms to outline that are central to the current understanding of the mechanisms of image scanning. The first is the island task that was initially conceived in the late 1970s whereby participants were required to memorise a map of an island with seven landmarks (Kosslyn et al., 1978). A list of pairs of landmarks were read out to the participant and the participant was required to imagine a black dot moving from the first to the second landmark. Results showed that RTs were longer between pairs of landmarks with greater distances between them. Such findings were met with criticism that participants are likely able to ascertain that the experimenter is expecting longer RTs between further distances (Pylyshyn, 1981). However, this claim is refuted by further experiments carefully designed to control for such confounds (Borst & Kosslyn, 2008; Borst et al., 2006; Borst & Kosslyn, 2010; Dror & Kosslyn, 1994). Specifically, an initial paradigm was derived whereby participants were required to memorise an array of dots, which subsequently disappeared. An arrow appeared in varying locations and participants had to determine whether the arrow appeared in the position of one of the dots or not. Here it was argued that participants had to scan a mental image of the dots to determine the correct answer and it was found that RT increased as a function of distance between the dots (Finke & Pinker, 1982). It might be argued that longer RTs in this paradigm could be explained by attention rather than MI in that recognition of targets is disrupted in crowded scenes (e.g., Strasburger, 2005). However, this finding has been demonstrated in a paradigm accounting for this confound. The revised paradigm involves presenting the outline of a grid with three filled in squares (see Figure 1.7). When the grid disappeared, an arrow appeared, and participants determined if the arrow pointed at a filled square or not. As expected, the further the distance between the arrow and the square on the grid, the longer the RT (Dror & Kosslyn, 1994). This was taken to confirm that individuals are in fact shifting attention across an image to determine the correct answer. Other experiments have extended these findings to demonstrate that the precision at which the spatial properties of dots held in mind could also be quantified by RT slopes (Borst & Kosslyn, 2010). Thus, collectively, these findings demonstrate that individuals represent varying distances in a mental image, can shift attention to inspect different elements of the image and that it is possible to capture this phenomenon using behavioural paradigms.

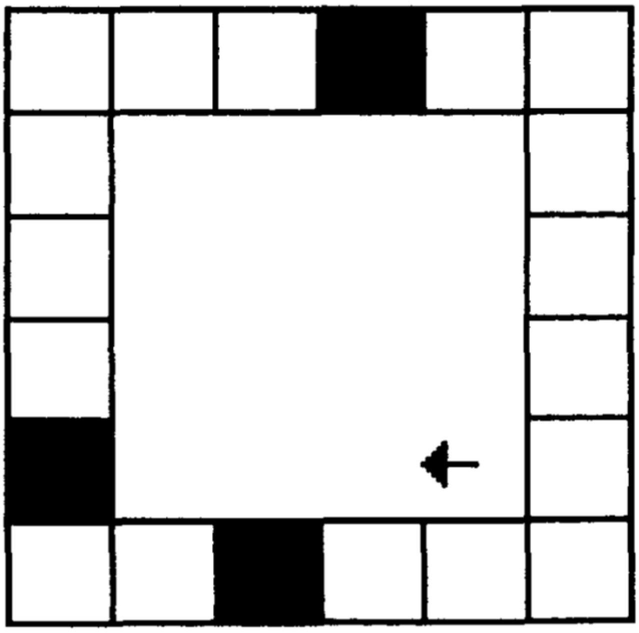


*Figure 1.7:* Image scanning paradigm whereby participants first memorise the grid and determine whether the subsequently presented arrow points towards a filled square or not. Grid and arrow shown for reference and not presented at the same time in the task. Reproduced with permission from Dror and Kosslyn (1994)

*1.2.3.1.8. Development of image scanning*

Studies investigating the development of image scanning throughout childhood have replicated the scanning effect observed in adults, however differences lie in the dependent variables derived and therefore the conclusions made regarding children's image scanning ability. Initially, Kosslyn et al.'s (1990) study derived a modified version of the task presented in Figure 1.7 involving the outline of a grid with filled squares. Participants viewed the grid outline with the filled squares. In contrast to the original paradigm, when the grid disappeared an 'X' or an 'O' were placed on the screen. If an X was placed, participants had to determine whether the 'X' overlapped a filled square, whereas if an 'O' was placed, participants had to determine whether the cell on the opposite side of the grid to the 'O' was filled. Thus, it was hypothesised that RTs would be longer in the 'O' trials compared to the 'X' trials, which would indicate that participants are required to scan in the 'O' trials. The expected RT difference was found. With respect to developmental progression, all age groups were slower, and accuracy was lower, in the trials where scanning was required compared to when it was not required.

There are just 2 other studies to date that have examined the developmental progression of image scanning in the context of a sub-component model of MI. The first to outline here adopted a very similar task to Kosslyn et al. (1990) but with more child-friendly stimuli whereby participants had to ask if the reference stimuli (an elephant) would catch the target stimulus (a ball) (Wimmer et al., 2017). In no scanning trials, the target stimulus was presented on the same grid as the reference and in scanning trials the target stimulus was presented in the opposite grid to the reference stimulus, as in Kosslyn et al.'s (1990) paradigm. Findings were replicated in RTs in that all age groups demonstrated longer RTs in the scanning trials compared to no scanning trials. Accuracy data revealed that only 6-, 8- and 10-year-olds showed higher accuracy in the no scanning trials compared to the scanning trials. Adults appeared to perform at ceiling, therefore paradigms sensitive to image scanning ability across age groups are required to establish developmental progression.

In addition, the use of such paradigms comparing a scanning or no scanning condition does not allow for the measurement of the precision in which individuals of different ages can shift attention across varying distances in mind. While the studies listed above denoted differences between scanning and no scanning trials in children, this does not identify whether children show the expected linear time-distance relationship. This was addressed in a recent modification of the island task. In this study, a map of an island with 5 landmarks and two signposts was presented (see Figure 1.8) (Wimmer et al., 2016). Signposts were included to assess how misleading top-down information impacted in image scanning. As in the original island task, researchers read a list of pairs of landmarks and participants were required to imagine a pirate parrot character walking from the first landmark in the pair to the second landmark. Findings showed that children aged 5-, 6- to 7-, 8- to 9- and 11-year-olds showed a significant increase in RT alongside increasing distance however 4-year-olds did not show this effect. Analysis of RT slopes showed the time-distance relationship observed in adults was present from age 5. Overall, this is the first study to investigate the developmental progression of the ability to shift attention across varying distances in mental images throughout childhood.

However, what remains unclear is how accurately individuals of different ages can represent distances in mental images.

*Figure 1.8:* Island scanning task: map of the island with landmarks placed at varying distances. Taken from Wimmer et al. (2016)

*1.2.3.1.9. The relationship between the components of MI throughout childhood and in adulthood*

Vital to our understanding of MI as a multi-faceted cognitive function is how the components of MI relate to one another throughout development. As outlined above, few studies have investigated the development of MI in the context of a sub-component model. In addition, the investigation of how the ability in each of the components relates to one another is limited. To date, there are just three studies that have investigated the development of components of MI and how they relate to one another in children, and these have yielded inconsistent results. Kosslyn et al. (1990) is the only study of MI in children to compare all four components in childhood and in adulthood. This demonstrated support for a separable-component model in that, with the exception of a positive relationship between error rates in mental rotation at age 8, the four components were statistically unrelated to one another for all age groups. More recent studies have compared pairs of components: image generation to image maintenance (Wimmer et al., 2015); and mental rotation to image scanning (Wimmer et al., 2017). While associations were not examined, Wimmer et al. (2015) found that participants demonstrated lower accuracy and slower RTs in image generation compared to image maintenance. This was argued to support initial suggestions that image generation involves additional processes, i.e., activating information from long-term memory, compared to image maintenance (Kosslyn et al., 2006; Kosslyn et al., 1990). Wimmer and colleagues (2017) also found no relationship between RTs in mental rotation and image scanning when age was controlled for. However, when examining age groups separately, 6-, 10-year-olds and adults showed significant positive relationships between mental rotation and image scanning RTs, which contrasts Kosslyn et al.'s (1990) finding. The authors speculated that developmental differences in working memory ability may underpin this finding, however this is yet to be tested directly. In this sense, the relationship between mental rotation and image scanning throughout childhood is currently unclear, and more clarity is required to determine whether image generation and image maintenance ability are related.

While the direct assessment of how each of the components of MI relate to one another is limited, some research in the MI literature has examined how MI abilities contribute to different types of reasoning and problem-solving. Firstly, in adults, relationships have been demonstrated between image scanning abilities and spatial transformation tasks, such as folding. Namely, image scanning RT slopes and accuracy were found to positively correlate with paper folding but did not correlate with scores on the object-imagery scale on the Object-Spatial Imagery Questionnaire (OSIQ) (Blajenkova et al., 2006; Borst & Kosslyn, 2010). The OSIQ includes items referring to the visual properties and precision of images, e.g., "My mental pictures are very precise representations of real things". Thus, this preliminary evidence suggests a distinction between ability in representing spatial properties in mental images compared to the visual precision of mental images. This notion is supported by recent studies of individuals with no visual imagery; recently described as individuals with Aphantasia (Zeman et al., 2015). While presenting with poor or absent ability to generate visual images, those with Aphantasia appear to score in the typical range on the spatial imagery scale of the OSIQ, which includes items such as "I can easily rotate 3D geometric figures" (Keogh & Pearson, 2017). Although limited to self-report data, these initial findings point towards a dissociation between the visual precision of images and spatial properties of images in adults.

### 1.2.4. Summary and open questions

The evidence outlined in this section demonstrates the need to fully examine MI components in the context of a sub-component model to distinguish how MI develops and to determine how each component of ability relates to one another. Current findings on the development of image generation are likely contaminated by task difficulty and the relationship between image maintenance and image generation remains unclear. Moreover, recent research has identified a suitable measure of image scanning for children, however it is currently unclear how accurately children represent distances in mind, despite findings with adults demonstrating differential performance when distance in image scanning is manipulated. In the adult literature, there is extensive research investigating the extent to which depictive representations are generated in MI. In the developmental literature, there is some investigation of how ability on sub-components of MI develop, however the extent to which visually depictive representations are generated in MI is not fully addressed. **Chapter 2** addresses these gaps in knowledge by modifying previous tasks to examine the extent to which a depictive theory of MI is supported in children aged 6-11 years and in adulthood and to characterise how each of the four components of the sub-component model develop and relate to one another.

## 1.3. ADHD, working memory and mental imagery

In an effort to determine a neuropsychological profile of ADHD, research has examined how ability in executive functions (EF), including inhibition (Alderson et al., 2007; Barkley, 1997), working memory (Karalunas et al., 2018; Rapport et al., 2008), and self-regulation (Sjöwall et al., 2013; Zelazo & Carlson, 2012) in ADHD compares to those abilities in typically developing (TD) children. Case-control designs, those that compare between-group differences between children with ADHD and TD children, have suggested that children with ADHD present with a range of EF deficits. Variability in how the combination of EF deficits presents in ADHD means that there is currently no single accepted theoretical model of neuropsychological attributes in ADHD (Barkley, 1997; Castellanos et al., 2006). Although the model of EF deficits is not clear cut, individuals with ADHD tend to present with working memory impairments that are characterised by poorer VWM (Gau et al., 2009; Gau & Chiang, 2013; Martinussen et al., 2005; Martinussen & Tannock, 2006; Nikolas & Nigg, 2013; Simone et al., 2015). That being said, the study of cognition in ADHD is beginning to acknowledge the vast heterogeneity in EF abilities, particularly in working memory (Bergwerff et al., 2019; Campez et al., 2020; Dajani et al., 2016; Fair et al., 2012; Kofler et al., 2019; Vaidya et al., 2020). The importance of understanding the development of working memory in ADHD is far-reaching given that working memory is positively associated with academic achievement (Best et al., 2011), and ADHD-related behaviours appear to mediate negative academic outcomes (Friedman et al., 2018; Orban et al., 2018). Despite the demonstrated association with academic achievement, training working memory in both typical populations (Sala & Gobet, 2017) and ADHD (Cortese et al., 2015; Rapport et al., 2013) has been found to have little effect on improving academic outcomes. Therefore, to support children in learning, a better understanding of the mechanisms of working memory in both typical development and ADHD is required. While there is a suggestion of shared mechanisms between MI and VWM in adulthood (Albers et al., 2013; Keogh & Pearson, 2011, 2014), this has not been investigated in typical development or ADHD. Therefore, examining the relationship between MI and VWM in typical development and ADHD could provide important insights into the mechanisms underpinning individual differences in WM impairments in ADHD.

### 1.3.1. Working memory deficits in children with ADHD: evidence for heterogeneity

Research has identified working memory deficits in ADHD, particularly in the visual domain[1]. Many studies have adopted Baddeley's multi-component model of working memory, which involves components accounting for the maintenance and manipulation of visual-spatial information (Baddeley, 2003). This has typically resulted in the use of forward and backward spatial span measures as an assessment of maintenance and manipulation in VWM respectively (Martinussen et al., 2005; Martinussen & Tannock, 2006; Simone et al., 2015). Findings have suggested a greater impairment in manipulation compared to maintenance measures of VWM (Martinussen & Tannock, 2006; Nikolas & Nigg, 2013; Simone et al., 2015), thus alluding to the suggestion that executive attention processes might underly VWM deficits (Kasper et al., 2012).

If core attention deficits underly VWM impairments, then we would expect homogenous impairments in VWM across individuals diagnosed with attention deficits, however this is not the case. Failure to determine a single, neuropsychological profile of ADHD over several decades of research employing case-control designs has led researchers to examine individual differences using data-driven analyses (Bergwerff et al., 2019; Campez et al., 2020; Dajani et al., 2016; Fair et al., 2012; Kofler et al., 2019; Vaidya et al., 2020). This has highlighted broad individual differences in VWM abilities in not only children with ADHD, but also typically developing children. Data-driven analysis techniques are applied to examine how variation in performance on each of the measures contributes to profiles of ability regardless of diagnostic characteristics. One such study found a between-group comparison of typically developing children and children with ADHD showed the expected effect; worse performance on all EF measures, including VWM, in children with ADHD. However, data-driven analysis in the TD group revealed distinct profiles, including a profile of poor performance in working memory relative to the other profiles derived. Distinct profiles were also revealed in the ADHD groups, and

[1] Note that developmental literature tends to use the term visuo-spatial working memory. However, for consistency with the evidence discussed from the adult literature, I will continue to use the term VWM (abbreviation of visual working memory).

this could not be explained by ADHD symptoms (Fair et al., 2012). Such findings are important for two reasons. Firstly, they demonstrate that between-group comparisons between children with ADHD and TD children can mask potential underlying variability in both groups, and secondly, they demonstrate that profiles of ability do not predict diagnostic groups. This has since been found in relation to general EF abilities (including VWM measures) by findings that have shown data-driven profiles of ability include both children with ADHD and TD children (Bergwerff et al., 2019; Dajani et al., 2016; Kofler et al., 2019). In these studies, only 62% of the ADHD sample (N=55, total N=136) presented with impaired working memory (Kofler et al., 2019) and 34% of the ADHD sample (N=93, total N=321) in another study were classified into the average EF ability profile (Dajani et al., 2016). Furthermore, a recent study assessing profiles of VWM found that children with ADHD (N=38, total N=89) accounted for 28.5% of a moderate VWM ability group and 21.4% of a high VWM ability group (Campez et al., 2020). Overall, the individual differences approach to investigating neurocognitive profiles of ability has demonstrated it is not possible to conclude that children with ADHD present with a working memory deficit, therefore further investigation as to mechanisms underlying profiles of ability is necessary.

### 1.3.2. MI as an explanation for individual differences in VWM in typical and atypical development

The contribution of working memory ability to learning and academic achievement has been investigated extensively throughout development and has largely demonstrated a positive relationship between VWM abilities and academic skills, such as mathematics and reading comprehension (e.g. Berg, 2008; Bull, Espy, & Wiebe, 2008; Giofrè et al., 2013). For children with ADHD, VWM abilities in early childhood predict poor academic achievement in adolescence (Sjöwall et al., 2017), impairments in VWM are linked to poor problem-solving ability in mathematics (Friedman et al., 2018) and impairments in VWM mediate inattentive behaviour during classroom instructions in this group (Orban et al., 2018). Despite this, training VWM has proven ineffective in both ADHD and TD children thus far (Cortese et al., 2015; Rapport et al., 2013;

Sala & Gobet, 2017). This demonstrates that further understanding of the cognitive mechanisms underpinning individual differences in VWM is required.

In line with this, recent research has adopted a more fine-grained approach to examining the mechanisms of VWM, rather than comparing group performance on composite measures. This study adopted a change-detection task where participants were required to detect changes in previously presented shape arrays that either changed dependent on shape only or colour and shape. Findings suggested that the visual quality of representations was no different between children with ADHD and controls as evidenced by similar patterns of performance between shape-only and colour-shape conditions in both TD and ADHD. However there were specific impairments in the ability to update visual information in the retrieval component of the task (Ortega et al., 2020). This therefore suggests that specific components of VWM are more impaired than others in ADHD, which might provide a possible mechanistic explanation as to why some present with VWM impairments in composite VWM measures and some do not. To fully understand what contributes to VWM in ADHD, it would be valuable to examine how the distinct components of MI, i.e., generating, maintaining, manipulating, and inspecting visual mental images, might relate to components of VWM ability.

### 1.3.3. Summary and open questions

Despite decades of research investigating how working memory presents in ADHD, the evidence is not clear cut. Case-control designs have suggested that children with ADHD have poorer VWM than TD children, however the application of data-driven analyses has revealed extensive individual differences, even in typical development. Furthermore, even though VWM is positively associated with academic achievement in typical development and appears to present a barrier to academic achievement in ADHD, training VWM does not lead to gains in academic measures in either typical development or ADHD. To develop evidence-based support for children in learning, we must take a step back and establish the mechanisms underpinning individual differences in VWM in both typical and atypical development. Research in adults has suggested that greater MI abilities are associated with greater VWM abilities, however the relationship between MI and VWM has not been investigated in typical development or in children with ADHD. To address this gap in the literature, the relationship between MI and VWM in typical development and children with ADHD is fully examined in **Chapter 2** and **Chapter 3** respectively with both a case-control design and a data-driven individual differences approach to analysis.

## 1.4. Thesis outline

In **Chapter 2**, a novel battery of MI tasks that are sensitive to characterising individual differences in the ability to generate, maintain, transform, and inspect visually depictive representations is applied to fully examine MI abilities in childhood (age 6-11 years) and adulthood. **Chapter 2** will also investigate how sub-components of MI are related to one another and how they relate to maintenance and manipulation measures of VWM in primary aged children and adults. This will not only elucidate the relationship between MI and VWM throughout development but will inform the vital theoretical groundwork required to establish how MI presents in children with ADHD.

In **Chapter 3**, the battery of tasks developed in **Chapter 2** are applied to characterise MI abilities in each of the components of MI in children with ADHD. Between-group comparisons are made between children with ADHD and TD children from the **Chapter 2** sample to examine patterns of abilities and the developmental level of MI abilities in children with ADHD. Individual differences in transdiagnostic profiles of MI and VWM in both children with ADHD and TD children are investigated using latent profile analysis.

Previous research, and the research presented in **Chapters 2** and **3**, compare ability *between* MI and VWM and inferences are made as to how individual differences in MI impacts VWM ability. However, to fully understand the mechanisms of this relationship, it is vital to test how variability in MI impacts VWM *within* a VWM task. In **Chapter 4**, an EEG study is presented using a novel VWM paradigm to directly examine how MI is recruited *within* a VWM task. This study firstly examines how the visual precision and capacity of representations held in VWM, as indexed by CDA, is modulated by the type of subjective ratings of MI. Second, the metacognitive link between MI and VWM is assessed by considering how individual differences in subjective ratings of MI relate to behavioural and neural correlates of VWM.

# Chapter 2: Evidence for a separable-component model of MI and differential relationships between MI, VWM and attention in childhood and adulthood

## 2.1. Introduction

There are three key gaps in the literature to address: 1) to examine the extent to which a depictive theory of MI is supported in children, 2) to clarify whether there is evidence for a separable-component model of MI in childhood and adulthood, 3) to investigate whether ability in each of the components of MI relates to ability in VWM and attention in both children and adults. There is a wealth of literature examining evidence for depictive representations in adults (e.g. Dijkstra et al., 2019; Kosslyn & Thompson, 2003; Pearson & Kosslyn, 2015), however there is limited evidence as to whether depictive representations are present in childhood. Moreover, the focus of research in adults on the format of representations has neglected the theoretical attribute that MI is a multi-faceted cognitive function made up of distinct sub-abilities. This sub-component model of MI has been adopted in the few studies that have examined the development of MI (Kosslyn et al., 1990; Wimmer et al., 2015, 2016, 2017), however tasks employed do not examine individual differences in visually depictive mental images in each of the sub-components of MI. Moreover, there is some contention over whether components are related or not in development. Thus, to address the first two gaps in the literature, the first aim of this chapter is to assemble a novel battery of MI tasks to examine individual differences in the extent to which visually depictive mental images are recruited in each of the sub-components of MI: image generation, image maintenance, mental rotation and image scanning, throughout primary school years (age 6-11 years) and in adulthood (aim 1) and secondly to establish whether a separable component model of MI is supported (aim 2). Lastly, preliminary evidence in adults has suggested a positive association between MI and VWM (Keogh & Pearson, 2011; 2014). This relationship has not been examined directly in children; however, some have argued VWM abilities might underpin developmental progression of MI components (Wimmer et al., 2017). Moreover, children with ADHD are often characterised as possessing VWM impairments (Kasper et al., 2012), however recent evidence points towards extensive individual differences in VWM abilities in both TD children and children with ADHD (Campez et al., 2020). Alongside this, studies are yet to directly investigate the role of attention in each component of MI in childhood or adulthood. Clearly, to fully examine how MI is characterised in children with attention deficits, the theoretical framework of MI development with respect to the role of attention and VWM must be comprehensively addressed. Therefore, the third aim of this chapter is to examine the relationship between components of MI, VWM and attention in primary school children and adulthood. This will provide clarity on the theoretical perspective that MI is a multi-faceted function and has implications for examining how components of MI support other high-order cognitive functions and learning outcomes.

### 2.1.1. Evidence for visually depictive mental images in childhood

The studies conducted thus far to examine the developmental progression of MI components have varied with respect to the experimental paradigms and the age groups tested. Evidence for the development of image generation and image maintenance appear to be largely dependent on the tasks employed. Kosslyn et al.'s (1990) study involved an image generation task requiring participants to imagine how a presented lowercase letter cue (e.g., 'g') would appear in uppercase (e.g., 'G') on a grid. Two X's were marked on the grid and participants had to determine whether both X's would fall on the grid squares occupied by the uppercase letter (see **Chapter 1, Section 1.2.3.1.2.** for a detailed description of the task). The image maintenance task involved memorising a grid with randomly filled squares and subsequently determining if the two X's presented on a subsequent blank grid fell onto any of the previously filled grid squares. Findings suggested that adult-like ability in image generation was not present until age 14 years (children aged 5 years, 8 years and 14 years, and adults were tested; Kosslyn et al., 1990). However, it is important to note that the image generation task was reliant on knowledge of capital and lowercase letters, thus limiting the age range the task was appropriate for and leading to exclusion of the 5-year-old group from most analyses. Findings from the image maintenance task demonstrated that all age groups showed increased error rates when maintaining an image for 3000ms compared to 500ms, and overall error rates decreased with age up to 14 years.

The development of image generation and image maintenance was more recently investigated using age-appropriate paradigms. Here, Wimmer et al. (2015) devised paradigms whereby participants had to memorise the location of a known object (e.g., a flower) and were either required to hold it in mind for 500ms or 3000ms (image maintenance) or they were required to complete a 30 second distractor task (image generation) (described in detail in **Chapter 1 Section 1.2.3.1.2.**). Participants (4-, 6-, 8-, 10-year-olds and adults) were subsequently tested to determine whether the subsequent known object presented (e.g., a vase) was in the same location as the remembered object (e.g., a flower). The distractor task was implemented to encourage generation of a visual image from long-term memory. In contrast to Kosslyn et al. (1990), it was reported that image generation ability was present from age 4 years with adult levels reached around 10 years of age. Moreover, Wimmer et al. (2015) found no effect of time maintained in image maintenance and while analyses of overall accuracy showed image maintenance ability was present from the youngest age tested, an adult level of ability at was reached at around 8 years of age. While the tasks presented in Wimmer et al. (2015) are more suitable for younger children, these tasks involved remembering the locations of known objects, which likely tap into visuo-spatial imagery as opposed to being able to measure visual precision of mental images. Moreover, verbal strategies might be possible in a paradigm using known objects. For example, it is possible to verbally remember that a picture of a vase was left of centre when cued to discriminate between a memorised and target object. Therefore, tasks examining the ability to generate and maintain a visually depictive mental image, or the visual precision of a mental image, are required. Given the contrasting findings reported here regarding developmental progression, it is not possible based on the current evidence to make conclusions regarding the development of these two components of MI. To understand how components of MI develop throughout childhood, and to ultimately produce impactful findings with implications for classroom learning, it is necessary to investigate how image generation and image maintenance abilities develop throughout the primary school years (aged 6-11 years).

Despite mixed findings regarding the development of image generation and image maintenance, investigation of mental rotation and image scanning tasks have shown robust developmental effects. For mental rotation, there is strong evidence for a linear time-degree of rotation effect (RTs increase as a function of increasing degrees of rotation) from age 5 (Estes, 1998; Frick et al., 2013; Marmor, 1975; Wimmer et al., 2017) through to adulthood (Borst et al., 2011; Shepard & Metzler, 1971). Thus, demonstrating the ability to transform images from early childhood. The linear time-distance effect, i.e. increasing response times as a function of distance scanned across a mental image, was also found in children from age 5 years (Wimmer et al., 2016), which is consistent with adult literature (Borst & Kosslyn, 2010; Kosslyn et al., 1978). The recent development of this child-appropriate image scanning task provides another opportunity assess the extent to which visually depictive mental images are supported in childhood. In this image scanning task, children were required to memorise a map with landmarks in specific locations and hold this image in mind while shifting attention between landmarks (Wimmer et al., 2016). While this finding suggests that young children are able to shift their attention across a visual image, further research is needed to provide evidence for visually depictive mental images in this component of MI, i.e., how accurately do children represent distances in visual images, and how does this develop throughout childhood? Evidence from the visual perception literature for less accurate distance estimation in childhood compared to adulthood implies that children may be less accurate at representing distance in mind (e.g. Giovannini et al., 2009; Thurley & Schild, 2018), however this is yet to be tested directly and is vital in understanding the extent to which children's MI involves depictive representations.

### 2.1.2. Is MI unitary or multi-faceted?

As discussed in **Chapter 1,** initial studies adopting a sub-component model of MI have found support for distinct components in adults (Kosslyn et al., 1984; Kosslyn et al., 2006), however more recent investigations with individuals with Aphantasia imply distinctions between spatial and visual components of MI (e.g., Pounder et al., 2021). This is partially supported by evidence in developmental groups: Kosslyn et al. (1990) found evidence for a positive relationship between image scanning and mental rotation in 8-year-olds only, whereas Wimmer et al. (2017) found evidence for this relationship in 6-year-olds, 10-year-olds, and adults, but not 8-year-olds. Thus, while there is indication that spatial components are related to one another and are distinct from visual components, further investigation is required to make this conclusion. Moreover, Kosslyn et al.'s (1990) study is the only to assess the relationship between the four sub-components throughout development; more recent studies have assessed components in pairs, i.e., image generation and image maintenance (Wimmer et al., 2015) and mental rotation and image scanning (Wimmer et al., 2017). Wimmer et al. (2015) found accuracy was higher in image maintenance compared to image generation in childhood, which was argued to suggest image generation involves additional processes to image maintenance (Wimmer et al., 2015), and the association between the two components was not examined. Using tasks designed to capture how visually depictive representations are recruited in each of the components of MI in children aged 6-11 years, this chapter will provide important clarifications on how the components of MI relate to one another throughout development.

### 2.1.3. The relationship between MI and VWM

While previous research has speculated that working memory might underpin the developmental progression of sub-components of MI (Wimmer, 2015; 2017), this has not been tested directly. As discussed in **Chapter 1, Section 1.2.1.3**, studies have compared abilities on composite measures of MI and VWM in adults suggesting a positive relationship between MI and VWM (Keogh and Pearson, 2011; 2014). However, research is yet to investigate how components of MI abilities are associated with components of VWM abilities in children or adults. Within this investigation it is important to distinguish between maintenance and manipulation components of VWM given that research has shown that these VWM processes are distinct and develop differentially. Namely, studies have provided support for Baddeley's multi-component model of working memory in that there is evidence for distinct processing components (i.e., manipulation vs. maintenance) and storage components that are domain-specific (visual vs. verbal). While VWM maintenance and manipulation are positively correlated (Alloway et al., 2006; Alloway & Alloway, 2013; Miyake et al., 2001), the relationship between maintenance and manipulation abilities appears to develop differentially (Alloway et al., 2006; Mammarella et al., 2008) and they become more integrated in adulthood (Donolato et al., 2017). This is thought to be dependent on developing executive attention processes throughout early childhood, which plays a greater role in VWM manipulation compared to VWM maintenance (Alloway et al., 2006; Cowan et al., 2005). Given that a primary aim of this thesis is to elucidate the relationship between MI and VWM in children with and without attention deficits, it is therefore advantageous to include measures of both VWM maintenance and manipulation in investigating how the components of MI relate to VWM ability.

### 2.1.4. The relationship between MI and attention

Central to our understanding of how information is held and manipulated in mind in VWM is the consideration of the role of attention control in working memory models. It is argued that attention control modulates both storage and processing of visual information in VWM (Baddeley, 2003; Cowan, 2011, 2014; Cowan et al., 2005). However, the role of attention in MI is less clear cut. From a theoretical standpoint, studies considering the role of attention in MI have been designed to primarily demonstrate that MI is not merely an epiphenomenon of attention (Thompson et al., 2011). Preliminary studies investigated how varying the focus of attention in MI in adults was associated with increased activation in the intraparietal sulcus, which is implicated in attention control (Wojciulik & Kanwisher, 1999), and thus was interpreted to suggest a top-down role of attention control in MI (Ishai, 2002; Ishai, 2010). More recent studies examined visual target selection and found that stronger MI of a specific target colour modulated attention towards a target of the same colour in subsequent visual search (Cochrane et al., 2020; Moriya, 2018). In line with these findings, it might be expected that MI is positively associated with attention control. How attention control supports different sub-components of MI is yet to tested in children or adults. Of the four (broadly dissociated) sub-components of MI, attention is explicitly referred to in one component; image scanning (Kosslyn et al., 2006), however, image scanning tasks are not designed to explicitly measure attention control. Research is therefore required to directly examine how attention control is associated with abilities in each of the MI components in children and adults.

In developmental research, a common measure of attention control (mostly termed response inhibition or inhibitory control, however for completeness, the term attention control is used herein) is the Go/No-Go task whereby subjects are required to respond to one stimulus as quickly as possible while inhibiting their response to another stimulus. Performance on the Go/No-Go task has been found to improve throughout the primary school years (Johnstone et al., 2007). In children with ADHD, some theoretical standpoints argue that impaired attention control is a primary characteristic of ADHD (Barkley, 1997), whereas others have suggested a multi-pathway model of ADHD accounting for heterogeneity in a range of executive functions including reward and delay aversion, as well as attention control (Castellanos et al., 2006). Whether a principal defining feature or not, attention control deficits in children with ADHD compared to typically developing children are commonly found (Berwid et al., 2005) and attention control impairments are specifically implicated with reference to functional abnormalities in fronto-parietal networks in ADHD (Hart et al., 2013). However, much like in the working memory literature in ADHD, there is heterogeneity in the occurrence of attention control deficits (Nigg et al., 2005; Willcutt et al., 2005). Clearly, the relative presentation of both VWM and attention control deficits in ADHD is complex and further investigation to understand heterogeneity in these cognitive processes in ADHD is required. Therefore, to contribute to a theoretical framework of how MI develops in those with and without attention deficits, it is first vital to understand how MI and attention control interact in primary school children.

### 2.1.5. The current study

The current study was designed to investigate evidence for a separable-component model of MI, to examine evidence for visually depictive mental images within this model and to address how components of MI relate to VWM and attention control throughout development in the primary school years and in adulthood. To address this, a MI battery designed to be suitable from 6 years to adulthood was developed. While it is acknowledged that previous studies have found the presence of MI abilities from age 4, the current battery of tasks was too difficult for children under 6 years. This was because, in line with a key aim of this research, amendments were made to the paradigm to assess the visually depictive nature of mental images. Furthermore, in order to examine adult-like abilities and assess the extent to which abilities differentiate over time it was vital to use measures that were also sensitive to adult-levels ability.

The image generation and image maintenance tasks presented here were adapted from Wimmer et al. (2015). The paradigm follows the same sequence as Wimmer et al. (2015), however abstract shapes were used instead of known stimuli (e.g., flowers). Based on Wimmer et al.'s (2015) study, the current image generation task involved a 30 second distractor task to encourage the visual image to be generated from long-term memory, whereas the image maintenance task involved a 3 second delay to encourage the visual image to be maintained. While it has been argued that known stimuli are easier to visualise (e.g. Kail, Pellegrino, & Carter, 1980), the use of known stimuli likely encourages verbalising strategies, whereas verbalising is a less efficient strategy for abstract shapes. Therefore, to determine the extent to which visually depictive mental images are recruited in image generation and image maintenance in children, participants were required to generate and maintain images of abstract shapes.

To measure visually depictive mental images in the behavioural image generation and image maintenance tasks, measures of high visual precision and low visual precision were derived, which refer to the visual precision of which visual information is held in mind. As such, in the current study, participants are required to memorise an abstract shape and detect the location of a difference in the target shape. Participants must first distinguish whether there is a difference in the target shape, and if there is a difference, they then must locate the exact difference, which allows us to categorically distinguish between generating and maintaining visual images with high precision and low precision, respectively. It is hypothesised that faster response times (RTs) for high precision responses compared to low precision responses. This would demonstrate that in high precision responses, participants are referring to a visual image of high precision, whereas in low precision responses, participants struggle to detect the difference using a visual image and therefore recruit a less efficient, non-imagery strategy such as verbalising.

These novel tasks were paired with a mental rotation task (as used in Gilligan, Thomas, & Farran, 2019; adapted from Neuburger et al., 2011) and an image scanning task modified from Wimmer et al. (2016). Given the extensive literature examining mental rotation, it is expected that the robust finding that RT increases as a function of degree of rotation will be replicated, thus implying that individuals are rotating an image in mind in this task. Similarly, in image scanning, it is expected that RTs will increase as a function of distance scanned (as in Wimmer et al., 2016). To examine evidence for individual differences in visual precision of mental images in image scanning, how mentally represented distance corresponds to actual distance in image scanning trials and in perception control trials will be assessed. This will be investigated by deriving a measure of distance represented from RT. The logic for this measure is based on the same logic that has been paramount to research investigating mechanisms of mental rotation; namely, the established finding that RT increases as a function of distance rotated. Given that previous research has shown RT increases as a function of distance scanned in image scanning tasks, it is reasonable to assume that we can use response times to estimate participant's ability to represent varying distances in mind.  If visually precise mental images are recruited, it is predicted that individuals will show similar patterns of error in imaged distance compared to perceived distance.

The second key contribution of this chapter is to examine: a) the relationship between components of MI developmentally and in adulthood and b) the relationship between each component of MI, VWM and attention control in childhood and in adulthood. It is first predicted a relationship between visual components (i.e., image generation and image maintenance) and an association between the components involving transformations (mental rotation and image scanning). Secondly, while the relationship between components of MI and VWM are yet to be investigated in childhood, the preliminary evidence for a positive relationship between MI and VWM in adults is argued to be dependent on the recruitment of visual images in both functions (Keogh & Pearson, 2011; 2014), thus, a positive relationship is also predicted in childhood. The VWM tasks presented here comprise of a forward span VWM task and backward span VWM task. The choice to include both measures is in line with suggestions that forward span represent maintenance in VWM and backward span represents manipulation in VWM (e.g. Bull, Espy, & Wiebe, 2008; Martinussen & Tannock, 2006). Thus, by including these measures, we can predict the relationship between MI and VWM at a more fine-grained level for the first time. That is, if children are using MI strategies to support VWM performance, as suggested in the adult literature, positive relationships are expected at least between the ability to maintain a visual image (image maintenance) and the ability to maintain in VWM (forward span), and a relationship between the ability to transform an image (mental rotation) and manipulate in VWM (backward span). Finally, the relationship between components of MI and attention control is yet to be investigated. Given the importance of attention control in theoretical models of VWM and the suggestion of a positive relationship between MI and VWM, it is vital to clarify how attention control is related to components of MI.

## 2.2. Method

### 2.2.1. Participants

Adult participants were recruited via social media and SONA recruitment systems at Birkbeck, University of London. Adults recruited via the university system were rewarded course credit and all participants were entered into a prize draw to win a £50 Amazon voucher. Children were recruited from two primary schools and received a “Young Scientist” certificate for participation. Adults were recruited in an initial experiment to pilot the novel tasks included in the MI battery; a priori power analyses suggested a required sample size of 58 participants. Because the pilot was successful and no changes were required; for completeness and to avoid repetition, the results are reported here as one experiment with both children and adults. The sample size suggested by a priori power analyses for the current study was 88. Children were recruited as part of a larger project that required 92 primary school children aged 6-11 years. Therefore, the sample consisted of 58 adults and 92 children (total N = 150). One participant from the adult group was excluded from the correlation analysis due to multiple mouse control errors at the start of the backward VWM span task leading to the task stopping after only 4 trials. The final sample for the analyses reported below included 92 children and 57 adult participants (total sample: N = 149). Participant details are in Table 1.

Table 2.1

*Demographics for sample*

| Age group | N (female) | Mean (SD) age in years; months |
|---|---|---|
| 6- to 7-year-olds | 31 (20) | 6;11 (0;6) |
| 8- to 9-year-olds | 26 (16) | 8;6 (0;6) |
| 10- to 11-year-olds | 35 (14) | 10;7 (0;5) |
| Adults | 58 (39) | 26;11 (6;8) |

The study was approved by the University Ethics Committee and was conducted in accordance with the Declaration of Helsinki. All adults provided written informed consent. Parents of child participants provided written informed consent and subsequently the study information was explained to the child and each child gave written informed consent.

### 2.2.2. Materials and procedure

All tasks were comprised and presented on a 23-inch ASUS LCD monitor running with MATLAB v2018a (The MathWorks Inc, Natick, MA) and PsychToolbox v3.0.14. Adults participated in a single 90-minute testing session in a quiet testing booth at the university. Children were tested at school whereby the 90-minute session was split into three approximately 30-minute sessions to minimise impact on the school day and participant fatigue. The order of tasks completed by each participant was counterbalanced. Instructions for all tasks were read by the researcher and written test instructions were presented on screen at the start of the task.

2.2.2.1. Image generation and image maintenance

Stimuli consisted of 4-sided abstract shapes with either 1 or 2 curved lines and 3 or 2 straight lines. All stimuli covered an approximately equal surface area within a space of 200x180 pixels. Participants were seated at a viewing distance of 57cm, from which images subtended a visual angle of 5.1° (height) by 4.7° (width).

At the beginning of each trial in the image generation task, an abstract shape was presented for 15 seconds and participants were instructed to try and remember the details of the shape. When the shape disappeared, a checkerboard mask was flashed onto the full screen for 500ms to remove aftereffects. Participants were then instructed to turn to a paper and pencil letter cancellation task (Massonnié et al., 2020). The task formed the distractor task whereby participants were presented with a 16x20 grid of capital letters, evenly spaced on A4 paper. Participants were required to read each line from left to right and cross out the letters T and G (5 Gs and 5 Ts were interspersed randomly). After 30 seconds, a visual cue appeared on the screen and the researcher alerted the participant to turn back to the computer task. Next, they were required to think of the previous shape and identify whether the next shape to appear on the screen looked different from the reference shape. The target shape was presented, and participants were required to press the corresponding labelled letter key on the keyboard to answer either “YES” (the shape looked different) or “NO” (the shape did not look different). If the participant pressed the “YES” key, they then had to use the mouse to click on the part of the shape they thought was different, if the participant pressed the “NO” key, the task moved on to the next trial. Piloting with children showed that RTs in the image generation and image maintenance tasks were confounded by mouse control errors, therefore for the image generation and image maintenance tasks, if participants responded “YES” they then pointed with their finger to the exact location of the difference on the screen and the experimenter clicked with the mouse. This is consistent with previous methods showing no significant differences between experimenter and participant responses (Wimmer et al., 2015).

In all trials, reference and target shapes were presented in the same location. In “same” trials, the same shape was presented. In “difference” trials, the target shape presented was the reference shape with one alteration: either the concave or convex attribute of the curved line was increased or the angle between two straight lines was reduced/increased (see Figure 1A). To ensure that the trial presentation was not predictable: presentation of same and different trials was randomised per participant, the number of same and different trials were not the same (10 different trials and 5 same trials) and there were 15 different locations that each pair were presented in, therefore the location of the shape in a new trial was also not predictable.

Precision of visual images was derived from “difference” trials only with two categories: high precision response and low precision response. High precision responses are responses whereby participants correctly identified a difference and the exact location of the difference, low precision responses are the responses whereby participants identified a difference but did not correctly identify the location of the difference. Two dependent variables were measured: percentage of “YES” responses to difference trials (out of all possible responses) that fall into each precision category and mean RT for yes responses within each category. RTs were calculated by summing the RT to the “YES” button press and the RT for the location mouse click. To check that low precision responses were not due to a bias to answer “YES” and subsequently guess the location of the difference, we computed one sample t tests of the percentage of “YES” answers against 100%, in image generation ($M$ = 45.29, $SD$ = 13.80) and image maintenance ($M$ = 52.18, $SD$ = 13.94), respectively. The percentage of “YES” responses was significantly lower that 100% in both the image generation ($t(149)$ = -48.55, $p$ < .001, $d$ = -3.96) and image maintenance tasks ($t(149)$ = -42.02, $p$ < .001, $d$ = -3.43). Thus, we can assume that the precision categories allow us to distinguish between level of precision of visual images generated and maintained in these tasks.

An overall accuracy dependent variable was calculated to investigate progression of image generation and image maintenance ability throughout primary school years and to include in the correlation analyses. High precision trials were given a score of 2, low precision trials were given a score of 1, correct same trials were given a score of 1 and incorrect responses were scored 0. Thus, resulting in a maximum possible overall accuracy score of 25 (10 difference trials, 5 same trials). For example, maximum score of 25 would be achieved by 10 difference trials with score of 2 (high precision) and 5 same trials with score of 1.

The image maintenance task differed from the image generation task only in that when the shape disappeared, participants were instructed to hold the shape in mind and look at the fixation cross presented in the centre of the screen for 3000ms. Participants were then presented with the target shape and responded as above (see Figure 1B). Scoring was as in the image generation task.

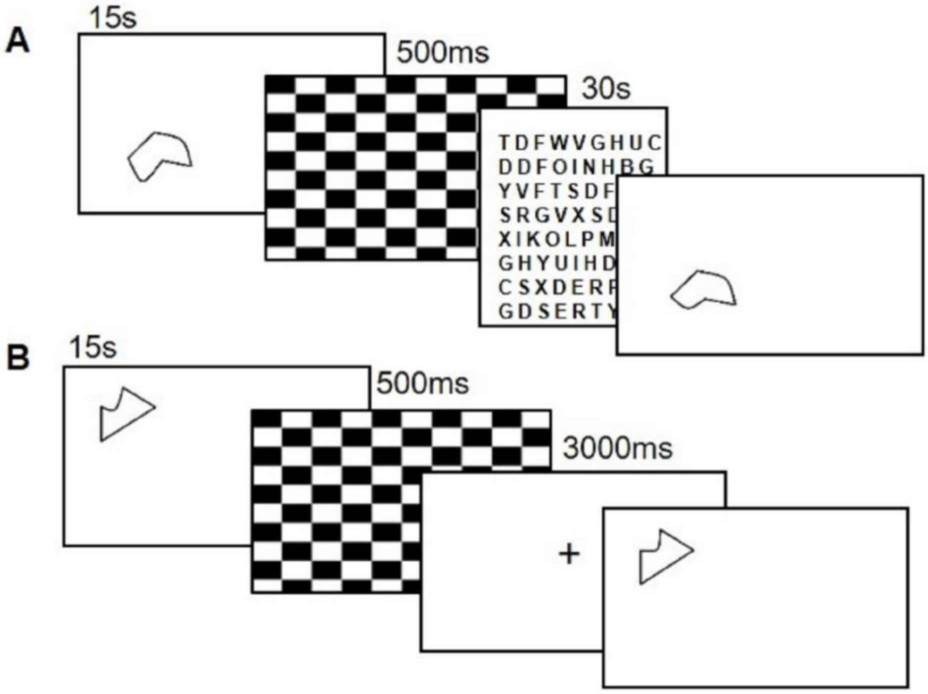



*Figure 2.1:* A) Image generation difference trial: presentation of the reference shape for 15s, presentation of checkerboard mask for 500ms, letter cancellation distractor task on A4 paper for 30s, presentation of test shape. B) Image maintenance difference trial: as above except instead of the letter cancellation task, a fixation cross appeared in the centre of the screen for 3000ms

2.2.2.2. Mental rotation

In the mental rotation task, participants were required to identify which of two mirror-imaged animal images located above a horizontal line, matched the target image below the line (Gilligan et al., 2019). The images above the line included a mirrored image of the target image and the target image at 0 degrees rotation (see Figure 2). Participants used the left arrow and right arrow keys on the computer keyboard to respond. Participants completed 4 practise trials at 0°, two with feedback. If participants scored less than 75% accuracy the practise was repeated. In experimental trials, the test image was presented at either 0°, 45°, 90°, 135° and 180° clockwise or anti-clockwise, 8 trials per degree of rotation summing to a total of 40 trials. The order of trial presentation was randomised per participant. RTs for correct trials were recorded.

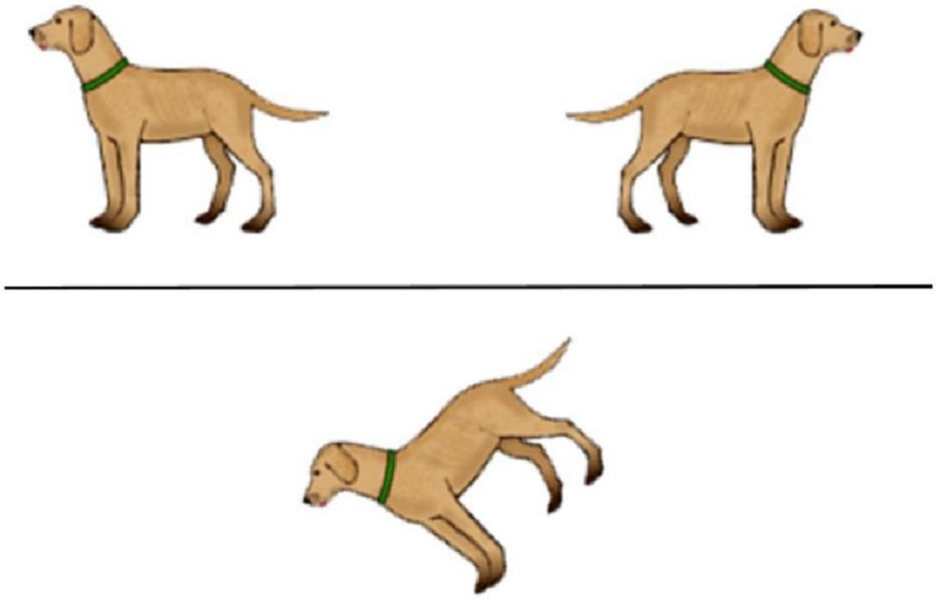


*Figure 2.2:* Example stimulus from the mental rotation task (45° anti-clockwise trial)

2.2.2.3. Image scanning

The image scanning task was adapted from Wimmer et al. (2016). On 4 practise trials, participants were shown a map of a park where they watched 'Percy the Pirate Parrot' walk between landmarks. Participants had to close their eyes and imagine the parrot walking between two specified landmarks after the experimenter said 'Start'. Participants were instructed to say 'Stop' when they imagined the parrot reaching the second landmark and were then asked how this compared to the Parrot walking across the screen and given further instruction if necessary.

Next participants viewed a map of an island containing a Lighthouse, Volcano, Hut, Pond and Tree for 45 seconds (see Figure 3). The signposts included in Wimmer et al.'s (2016) map of the island were removed as we did not include research questions regarding influence of conceptual knowledge. Following this the landmarks disappeared leaving an empty island. Participants recalled each landmark and then used the mouse to click on the location of each landmark, the landmark then appeared in the correct location and participants could compare their prediction. The island then disappeared, and the participants were instructed to close their eyes and imagine the map of the island. The experimenter read a list of 5 pairs one at a time in the following order: Lighthouse-Tree (262mm), Lighthouse-Volcano (81mm), Hut-Pond (100mm), Hut-Lighthouse (70mm), Lighthouse-Pond (154mm). As in the practise, participants were instructed to imagine the Parrot walking from the first landmark to the second landmark after the experimenter said ‘Start’ and for the participant to say ‘Stop’ when they imagined the parrot reaching the second landmark.

Finally, participants completed perception control trials whereby the map was visible on the screen while the list of pairs of landmarks was tested. This was to ensure participants were able to perceptually differentiate between the varying differences. Participants were instructed to trace their eyes between the landmarks where the parrot would be walking. For both MI and perceptual control conditions, RT was measured as the dependent variable. This was recorded by an experimenter click at ‘Start’ and an experimenter click at the participants’ ‘Stop’ command. The experimenter response method replicates that of previous experiments which have found no significant difference between experimenter and participant response (Wimmer et al., 2015; Wimmer et al., 2016).

To investigate the ability to accurately represent varying distances in mind, we compared imaged distance ratios to actual distance ratios using the following formula. Actual ratios were calculated by dividing the 2$^{nd}$, 3$^{rd}$, 4$^{th}$, and 5$^{th}$ actual distances by the first distance (e.g., actual ratio 1: 81mm/70mm = 1.16). Imaged distance ratios were then calculated in the same way, by dividing mean RT of the 2$^{nd}$, 3$^{rd}$, 4$^{th}$, and 5$^{th}$ distances by the mean RT of the first distance.


*Figure 2.3:* Island scanning task with distances: Hut-Lighthouse (70mm), Lighthouse-Volcano (81mm), Hut-Pond (100mm), Lighthouse-Pond (154mm), Lighthouse-Tree (262mm). Lines between landmarks and distance labels added for reference and not presented during task

2.2.2.4. VWM

The VWM task involved a backward span task (Morris et al., 2019) and a forward span task administered with the forward span task preceeding the backward span task, in line with standard procedures (Simone et al., 2015). Participants were presented with a 3x3 grid of lily pads (see Figure 4). To practise, participants completed 3 trials with a sequence of 2. If participants clicked outside of the lily pad or were incorrect, the practise ended and was repeated. In forward span experimental trials, participants watched a frog appear on multiple lily pads ranging from 2-9. They then were required to click on the lily pads first to last. Each sequence length was repeated 4 times, and sequence length increased by one in each new block. If participants made two or more errors in one block, the task stopped. The backward span task followed the same procedure including practise trials, except participants had to click on the lily pads back to front, i.e., in the reverse order. The dependent variable recorded from each task was the maximum sequence length reached (range 2-9).

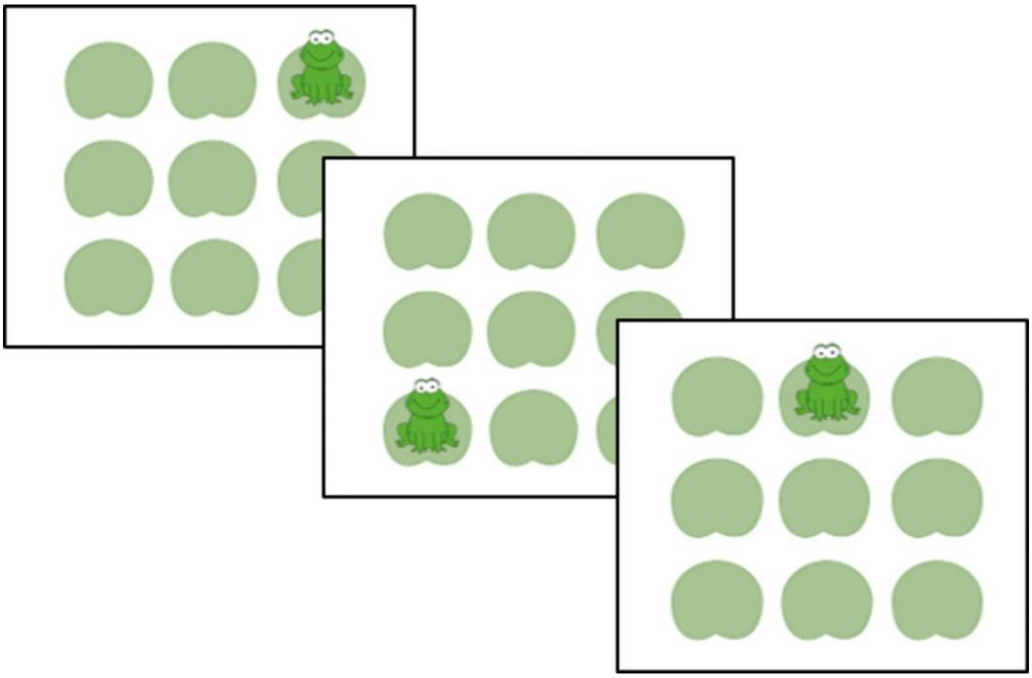


*Figure 2.4:* VWM – example sequence from 3 span trial

2.2.2.5. Go-No/Go

The Go-No/Go task is as used in Farran et al. (2015). A randomised series of 5-cm diameter red, orange, purple and yellow solid circles were presented on a white background. Participants were instructed to press the space bar as quickly as they could as soon as they saw a coloured circle appear on the screen, unless the circle was red. They were instructed that if the circle was red, they were to refrain from pressing the button and wait for the next circle to appear. If the participant pressed on a red circle, an 'error' noise was heard, and the circle disappeared. If the participant did not respond, the circle remained on the screen for 2 seconds before moving onto the next trial. Red trials were presented on 25% of all trials and trial presentation was randomised. There were 8 practise trials and 64 experimental trials. Error rates were initially chosen as the dependent variable; however, performance was at ceiling in the child and adult samples, therefore, mean RT from correct answers formed the dependent variable for attention control.

### 2.2.3. Analysis strategy

Analyses were preregistered (https://osf.io/rk7f5/). All analyses reported in this chapter are detailed in the pre-registration document and are detailed alongside each task analysis below. The only modifications made were to exclude the variable of RT for incorrect trials from the image generation and image maintenance tasks and the Go-No/Go measure of RT variability is not reported here. The incorrect RT variable was not included because only 41% of children and 38% of adults incorrectly identified a same trial as different, therefore there was not sufficient data for this analysis. To correct for multiple comparisons in the correlation analyses, $p$ values were Bonferroni corrected (alpha level 0.05 / 21 comparisons) resulting in $p = .002$. The preregistration includes a correlation matrix per age group based on Wimmer et al.'s (2017) effect sizes. However, when this was carried out the effect sizes were very small (mean r for 6-7-year-olds: -.014; 8-9-year-olds: .012; 10-11-year-olds: .023). Sensitivity power analyses allow for the estimation of a range of possible effect sizes dependent on alpha level, power, and sample size (Lakens, 2021). Plotting a sensitivity curve revealed as many as 85 participants would be needed to detect an effect size of .400 with alpha level set to .002 (multiple comparisons cutoff) and power at 0.8 (see Figure A.1 in **Appendix**). Therefore, a much larger sample per age group would be required to be sufficiently powered. Instead, one correlation matrix was conducted controlling for age in years and collapsed across the three groups of primary aged children and a second correlation matrix was conducted with adults. On account of non-normal data, non-parametric Spearman correlation matrices were conducted. Analyses of slopes of best fitting lines on mental rotation and image scanning RTs are additional to the preregistration, as well as the analysis of variance (ANOVA) on perception control RT data.

Tests of normality revealed that most variables were not normally distributed (Kolmogorov-Smirnov Test, $p$s < .05). Sample sizes per age group are adequate for use of parametric analyses in line with the central limit theorem (Field, 2013). As there are unequal numbers of participants across groups, Welch's $F$ is reported for one-way ANOVAs and Games-Howell correction is applied to post-hoc comparisons. Within-subject pairwise comparisons are reported with Bonferroni corrections. As we are comparing adult and child samples in the same analyses, some variance ratios for RT variables exceeded that appropriate for assumptions of parametric tests. Levene's test of equality of variance was violated for the RT dependent variables in all tasks. Given that the sample sizes in each age group are unequal, RT variables were log transformed to meet the assumption of normality in multivariate tests. Levene's tests were no longer significant following log transformations of RTs ($p$s > .05). There were no differences between the results from the transformed vs non-transformed data.

Cronbach's alpha was conducted to assess the internal consistency of novel measures: image generation, image maintenance and the image scanning imaged distance ratios. The structure of the scores for items in the image generation and image maintenance tasks (i.e., possible scores were 0, 1, 2) results in low standard deviations per item on account of the small range of possible scores. Because of this, interitem covariance is not meaningful and it is not appropriate to conduct Cronbach's alpha reliability analyse (Cronbach's alpha uses mean interitem covariance) on the image generation and image maintenance accuracy measures (Cortina, 1993). Thus, for image generation and image maintenance tasks, internal consistency is calculated using RT. This revealed good internal consistency for both image generation ($\alpha$ = .70, SD = .70) and image maintenance ($\alpha$ = .70, SD = .71). Finally, good internal consistency was also revealed for the image scanning imaged distance ratios ($\alpha$ = .70, SD = .78). With regard to the reliability of the previously used measures, the VWM measures have been used in previous research with over 300 children and found to have good test-retest reliability (Morris et al., 2019). The mental rotation task has also been used in a number of previous studies (Gilligan et al., 2019; Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011) and is a well-established paradigm.

## 2.3. Results

### 2.3.1. Image generation

A one-way ANOVA examining age group differences in overall accuracy in image generation showed a significant main effect of age (*Welch's F*(3,70.51) = 6.86, $p$ < .001, $\eta^2$ = .12), whereby post-hoc comparisons show a significant improvement between children 6- to 7-year-olds (*M* = 13.00, *SD* = 3.25) and 10- to 11-year-olds (*M* = 15.97, *SD* = 3.19; *p* = .002) and between 6- to 7-year-olds and adults only (*M* = 15.97, *SD* = 3.52; *p*<.01), but no difference across groups from age 8-9 years (*M* = 14.19, *SD* = 3.20) onwards (*p*s > .05 for all).

A mixed ANOVA of precision response type (logRTs) with a within-participant factor of precision (high, low) and a between-participant factor of age group (6-7 years, 8-9 years, 10-11 years, adults) demonstrated all participants responded significantly more quickly when responses were of high precision compared to responses of low precision ($F(1,146) = 54.04$, $p < .001$, $\eta_p^2 = .27$, see Figure 2.5.C). There was no main effect of age and no interaction with age ($F < 1$ for both). This therefore supports the hypothesis that in high precision responses, individuals are referring to a highly visual mental image, whereas in low precision responses, participants are likely searching for an alternate less efficient strategy on account of low precision responses. Mean and SDs of raw RTs in milliseconds for all RT variables can be found in the **Appendix**.

### 2.3.2. Image maintenance

A one-way ANOVA investigating age group differences in overall accuracy in image maintenance revealed a significant main effect (*Welch's* $F(3,66.07) = 11.49$, $p < .001$, $\eta^2 = .21$), whereby there were significant improvements between 6- to 7-year-olds (*M* = 14.35, *SD* = 3.61) and the oldest children (10- to 11-year-olds: *M* = 16.54, *SD* = 2.67; *p* = .037) and 6- to 7-year-olds and adults (*M* = 17.91, *SD* = 2.57; *p* < .001). There was also a significant improvement between 8- to 9-year-olds (*M* = 14.89, *SD* = 3.08) and adults (*p* < .001; all other *ps* > .05).

A mixed ANOVA of precision response type (logRTs) by age group, revealed a significant main effect of precision ($F(1,146) = 61.65$, $p < .001$, $\eta_p^2 = .29$), which indicated that participants showed significantly faster RTs in high precision responses compared to low precision responses (see Figure 2.5.D). There was a significant main effect of age group ($F(3,146) = 32.52$, $p < .001$, $\eta_p^2 = .40$); 6- to 7-year-olds were significantly slower to respond than all other age groups (*ps* < .05), 8- to 9-year-olds were slower than adults (*p* < .001) and 10- to 11-year-olds were slower than adults (*p* < .001). There was no interaction between precision and age ($F(3,146) = 1.84$, $p = .143$, $\eta_p^2 = .04$). Taken together, this supports the hypothesis that individuals are able to maintain mental images of high precision as characterised by faster responses, compared to searching for an alternate, less efficient strategy in low precision as characterised by slower responses.

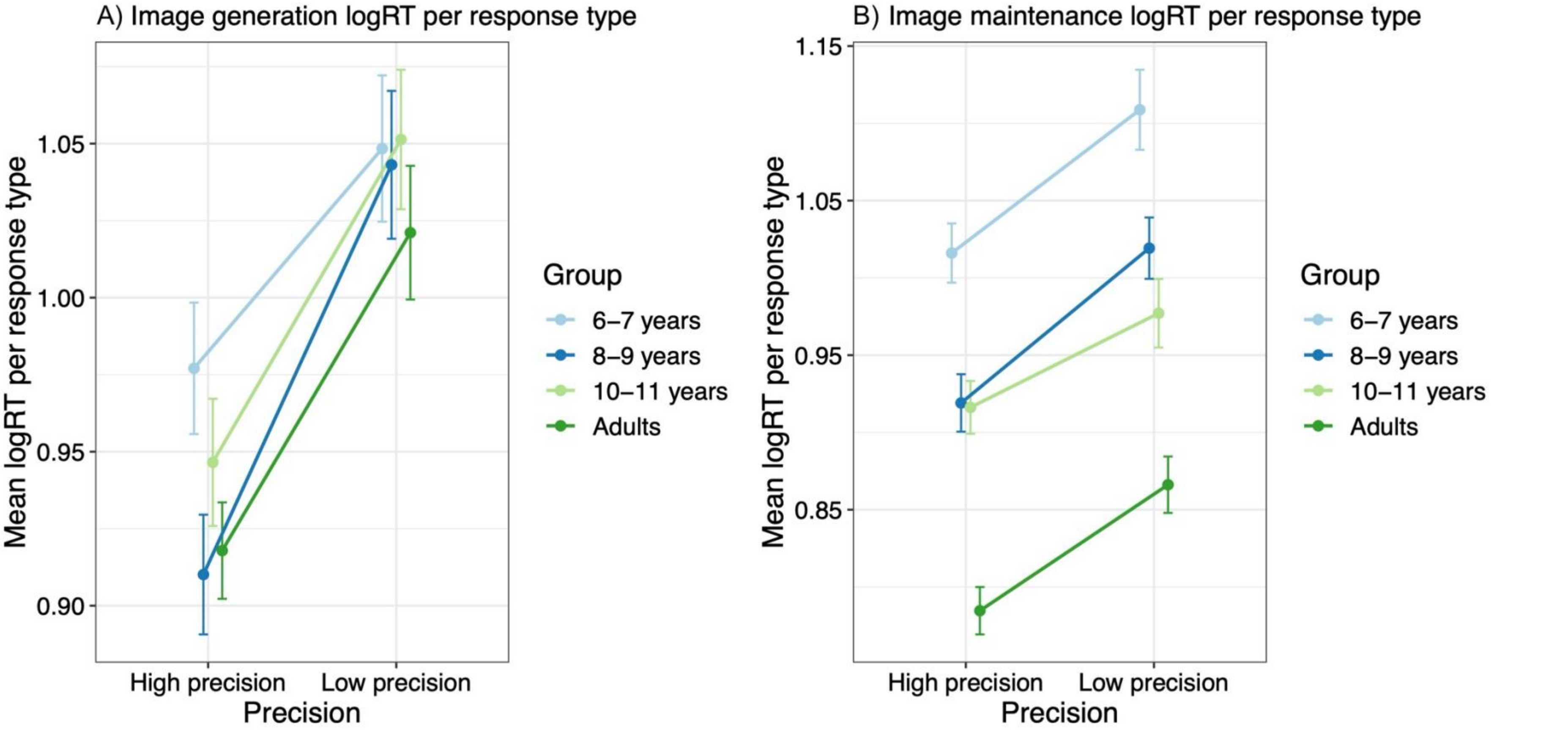



*Figure 2.5:* Mean and standard error (SE) logRTs in high and low precision responses per age group in image generation (A) and image maintenance (B)

### 2.3.3. Image generation vs. image maintenance

To investigate the relationship between performance in image generation and image maintenance, a mixed ANOVA was conducted on task (image generation, image maintenance) by age group for high precision trials only, whereby the dependent variable was logRT, respectively. The ANOVA revealed a significant main effect of task with a small effect size ($F(1,146) = 6.98$, $p = .009$, $\eta_p^2 = .05$), which suggested faster RT in image maintenance compared to image generation. A significant main effect of age ($F(3,146) = 17.66$, $p < .001$, $\eta_p^2 = .26$) showed faster RTs between all older age groups and younger age groups (*ps* < .05), except for between 10- to 11-year-olds and 8- to 9-year-olds ($p > .05$). There was also a significant interaction with age ($F(3,146) = 15.22$, $p < .001$, $\eta_p^2 = .23$). Simple effects analysis showed a significant main effect only for adults, which showed faster RTs in image maintenance ($M = .92$, $SD = .09$) compared to image generation ($M = .79$, $SD = .12$; $p < .001$, all other *ps* > .05).

### 2.3.4. Mental rotation

2.3.4.1. Mental rotation RT

A mixed ANOVA was conducted on the mean logRT of correct trials per degree of rotation (0°, 45°, 90°, 135° and 180°) by age group. There was a significant main effect of age ($F(3,146) = 69.54$, $p < .001$, $\eta_p^2 = .58$) on account of significantly longer RTs between each age group and adults, as well as longer RTs in 6- to 7-year-olds compared to 10- to 11-year-olds and in 8- to 9-year-olds compared to 10- to 11-year-olds (all *ps* < .001). There was a main effect of degree of rotation, which was best explained by a significant linear contrast ($F(1,146) = 282.05$, $p < .001$, $\eta_p^2 = .66$) demonstrating that as degree of rotation increased, logRT increased. There was also a significant interaction between degree of rotation and age group ($F(9.97,485.04) = 1.92$, $p = .05$, $\eta_p^2 = .04$). However, follow up analyses revealed significant linear contrasts best explained the linear effect of RT for all age groups (6- to 7-year-olds: $F(1,30) = 23.42$, $p < .001$, $\eta_p^2 = .27$; 8- to 9-year-olds: $F(1,25) = 32.59$, $p < .001$, $\eta_p^2 = .57$; 10- to 11-year-olds: $F(1,34) = 138.47$, $p < .001$, $\eta_p^2 = .80$; adults: $F(1,57) =$

218.14, $p < .001$, $\eta_p^2 = .79$). Means and standard deviations of raw RTs can be found in the **Appendix** (Table A.1.2.2).

Lines of best-fit were calculated to determine the slope value, i.e. the gradient of the line, of mean correct RT across the five degrees of rotation (0°, 45°, 90°, 135°, 180°) for each participant. A one-way ANOVA of slope values by age group revealed a significant main effect of age group (*Welch's* $F(3,24.88) = 5.52$, $p = .002$, $\eta_p^2 = .08$). Post hoc comparisons revealed a significantly steeper slope between 8- to 9-year-olds ($M = .64$, $SD = .61$) and adults ($M = 0.31$, $SD = .23$) ($p = .005$). There were no other between group differences (6- to 7-year-olds: $M = .43$, $SD = .52$; 10- to 11-year-olds: $M = .53$, $SD = .36$; all $p$s > .05).

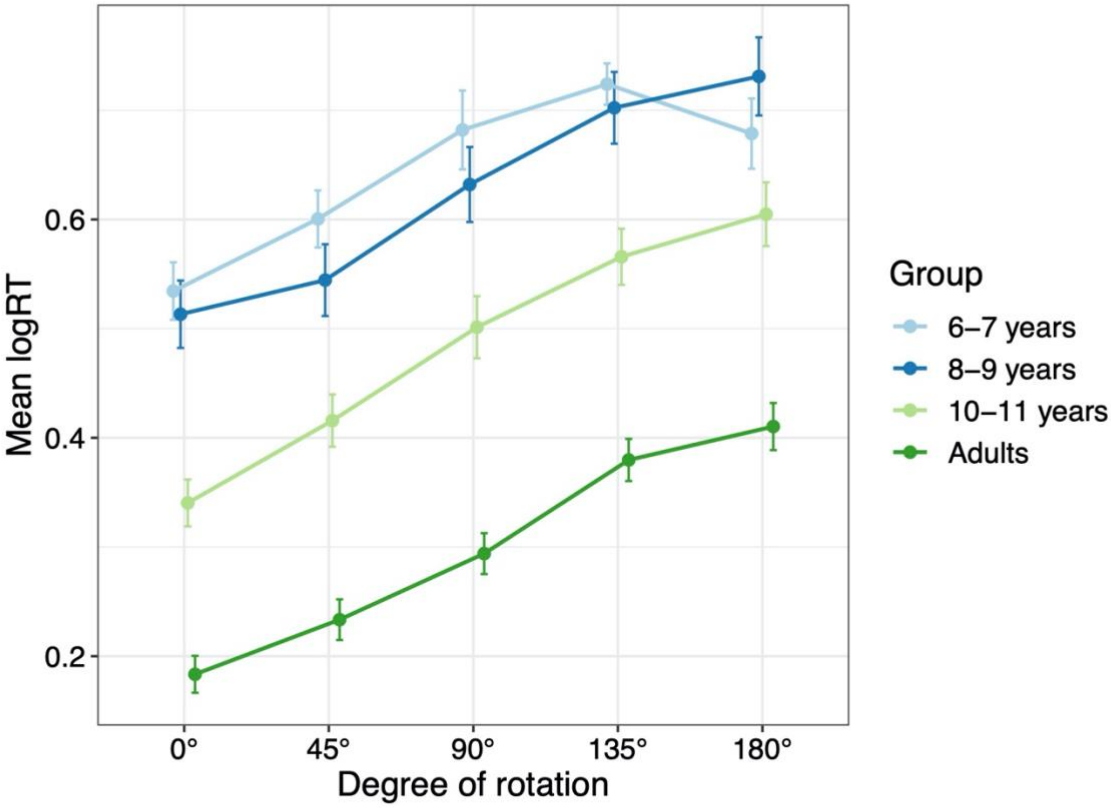



*Figure 2.6:* Mean and SE logRT per degree of rotation for each age group in the mental rotation task

2.3.4.2. Mental rotation accuracy

An equivalent ANOVA was conducted on percentage accuracy by age group. There was a significant main effect suggesting as degree of rotation increased, accuracy decreased ($F(1.95,284.69) = 68.44$, $p < .001$, $\eta_p^2 = .32$). There also a significant main effect of age ($F(3,146) = 13.71$, $p < .001$, $\eta_p^2 <$

.22), whereby comparisons indicated significant improvements in accuracy between all age groups (*ps* $< .05$), apart from between 6- to 7-year-olds and 8- to 9-year-olds ($p > .05$) and between 10- to 11-year-olds and adults ($p > .05$). There was no interaction with age ($F < 2$, $p > .05$).

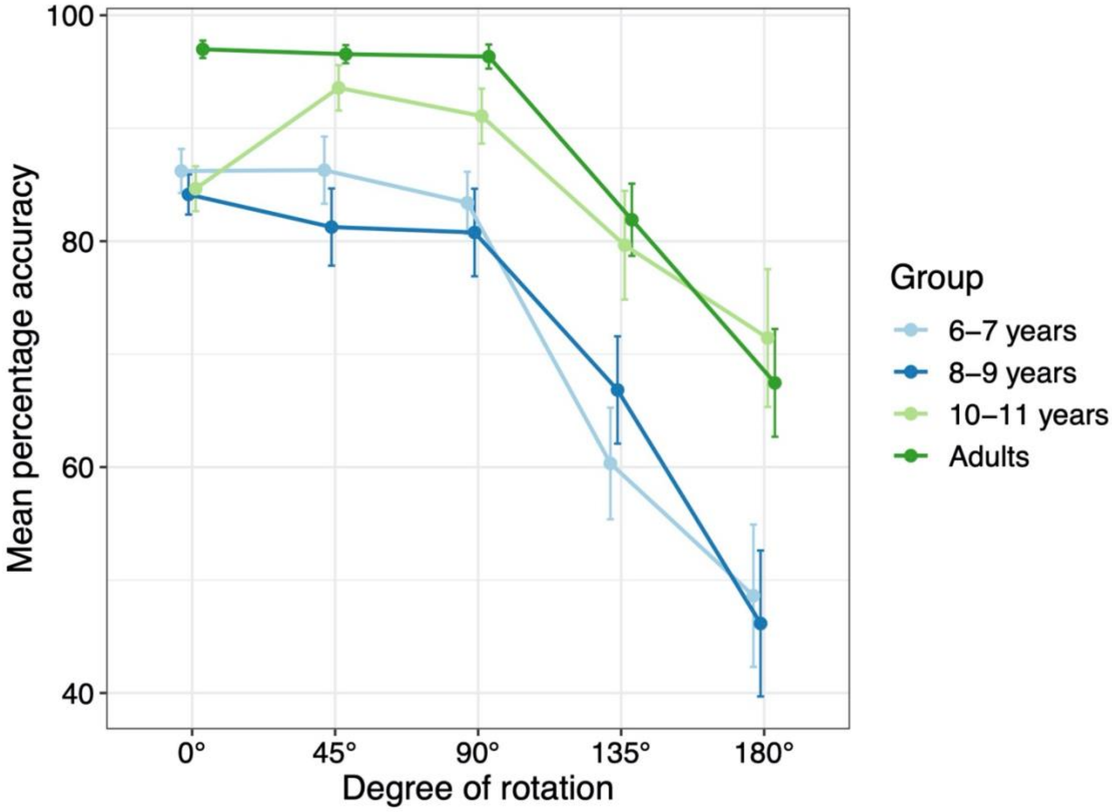



*Figure 2.7:* Mean and SE percentage accuracy per degree of rotation in mental rotation for each age group.

### 2.3.5. Image scanning

A mixed ANOVA of mean logRT with factors of distance (70mm, 81mm, 100mm, 154mm, 262mm) and age group revealed a significant main effect of distance ($F(3.36,490.06) = 116.57$, $p < .001$, $\eta_p^2 = .44$) showing that as distance increased, logRT increased (see Figure 2.8). Pairwise comparisons revealed no significant differences between 70mm, 81mm and 100mm (*ps*>.05). This might be expected as the 3 shortest distances are very comparatively close together with differences of only 11mm-19mm between the landmarks. In turn, there were significant increases in RT between each smaller distance and the two largest distances, as well as a significant increase between the 2nd largest and largest distance (all *ps* $< .001$). A significant main effect of age group ($F(3,146) = 2.751$, $p < .05$, $\eta_p^2 = .05$) indicated that 10- to 11-year-olds had significantly longer responses than adults ($p < .05$; all other comparisons: $p > .05$). There was a significant interaction between distance and age group ($F(10.07,490.06) = 4.20$, $p < .001$, $\eta_p^2 = .08$). Follow-up analyses showed a main effect of distance in all age groups (6-7-years: $F(3.06,91.66) = 16.43$, $p < .001$, $\eta_p^2 = .35$; 8-9-years: $F(4,100) = 8.07$, $p < .001$, $\eta_p^2 = .24$; 10-11-years: $F(3.58,121.64) = 42.78$, $p < .001$, $\eta_p^2 = .56$; Adults: $F(4,228) = 135.12$, $p = .001$, $\eta_p^2 = .70$). Pairwise comparisons showed only adults had the exact pattern of RT differences outlined above (all *p*s $< .001$). Children age 10-11 years showed almost the same pattern as adults, however there was no significant increase in RT between the 2nd furthest distance and the furthest distance (154mm < 262mm; $p = .116$). Children age 6-7 and 8-9 years showed significant increases between the two shortest distances and two furthest distances only (*ps* $< .05$). Means and standard deviations of raw RTs can be found in the **Appendix** (Table A.1.2.3).

As in mental rotation, slope values were calculated for each participant on non-transformed RT and a one-way ANOVA examining age group differences in slopes was conducted. There was no main effect of age group (*Welch's* $F(3,62.07) = 1.65$, $p = .186$, $\eta_p^2 = .02$).

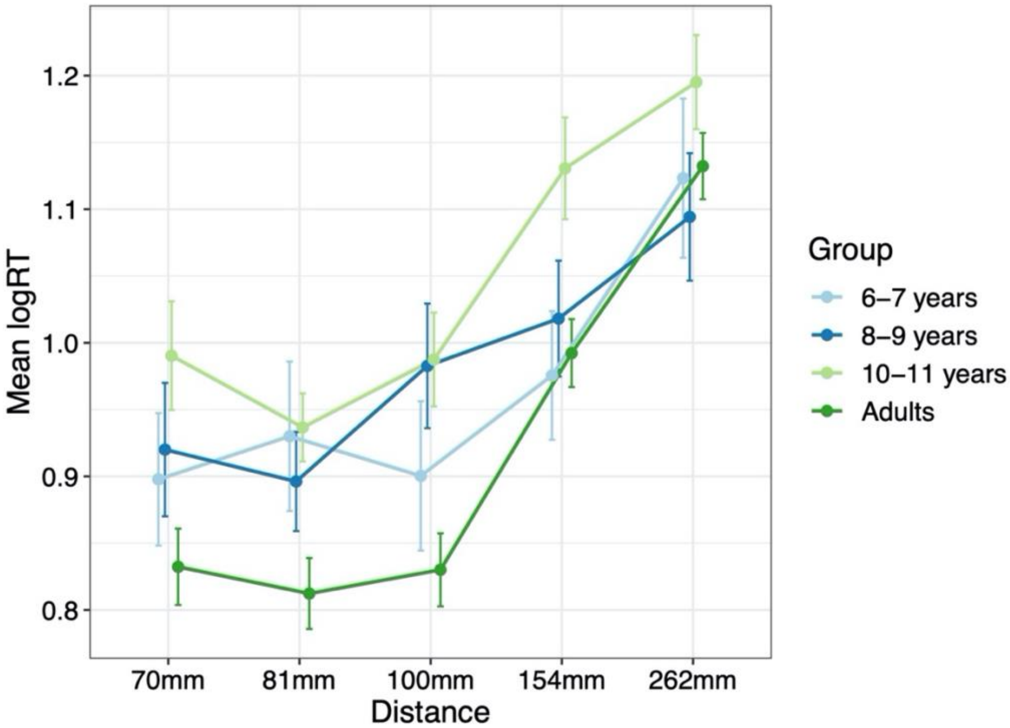



*Figure 2.8:* Mean and SE logRT per distance for each age group in the image scanning task

The equivalent mixed ANOVA was conducted on the perception control trial logRT with distance and age group. This showed broadly the same pattern as the image scanning condition. There was a significant main effect of distance, ($F(2.41,352.43) = 192.25$, $p < .001$, $\eta_p^2 = .57$), showing as distance increased, RT increased. Pairwise comparisons revealed significant differences between all comparisons ($p < .05$), except for the difference between 70mm and 81mm ($p > .05$) and between 81mm and 100mm ($p > .05$). There were significant increases in RT between all other short distances and longer distances (*ps* < .001). While, there was no significant effect of age group ($p > .05$), there was a significant interaction between perception logRT and age group ($F(7.24,352.43) = 35.99$, $p = .02$, $\eta_p^2 = .05$). Follow up repeated measures ANOVAs revealed each age group demonstrated a significant main effect of distance (6- to 7-year-olds: $F(2.19,65.98) = 28.49$, $p < .001$, $\eta_p^2 = .49$; 8- to 9-year-olds: $F(4,100) = 19.69$, $p < .001$, $\eta_p^2 = .44$; 10- to 11-year-olds: $F(2.29, 78.08) = 36.86$, $p < .001$, $\eta_p^2 = .52$). Pairwise comparisons showed 6- to 7-year-olds and adults showed the same effect as above (*ps* < .001), whereas 8-9 years and 10-11 years showed the same effect except no significant increase in RT between 154mm and 262mm (*ps* > .05). Slopes of best fitting lines on RT were calculated per participant and a one-way ANOVA was conducted on slopes by age group to investigate whether steepness of slope was different between age groups; there was no significant main effect of age group (*Welch's* $F < 1$), in line with the ANOVA examining image scanning slope values. Means and standard deviations of raw RTs can be found in the **Appendix** (Table A.1.2.4).

#### 2.3.5.1. Evaluating the ability to internally represent distance in image scanning

The relationship between actual distance ratios (ratio 1: 81mm/70mm = 1.16, ratio 2: 100mm/70mm = 1.43, ratio 3: 154mm/70mm = 2.2, ratio 4: 262mm/154mm = 3.74) and mean imaged distance ratios (derived from RT) per age group are displayed graphically in Figure 2.9.

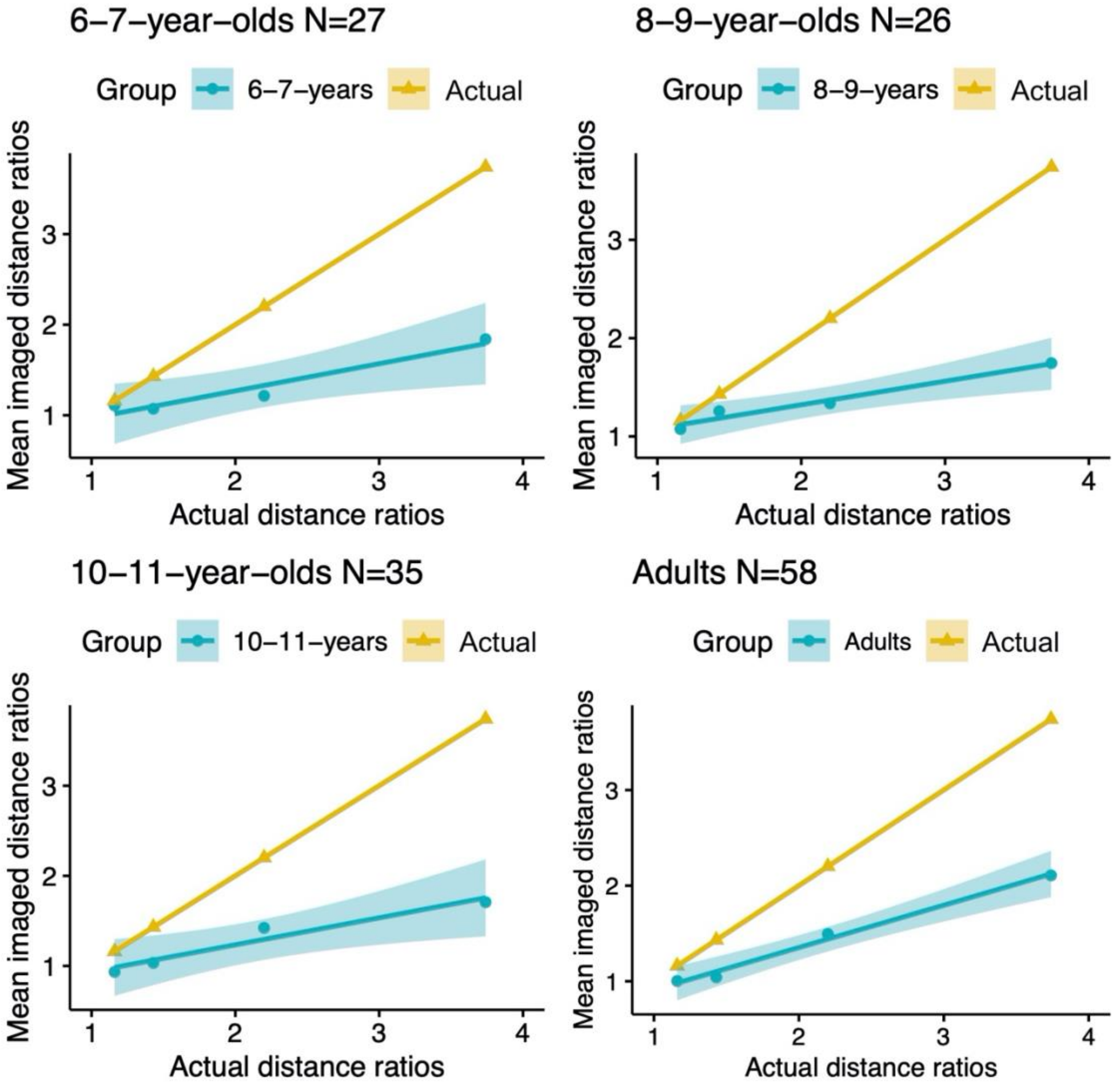



*Figure 2.9:* Scatterplots depicting the relationship between actual distance ratios and mean imaged distance ratios with confidence intervals and best fitting regression lines. Actual distance ratio line added for reference

Visual inspection suggests that the ratios are highly linearly related. That is, as actual distance increases, imaged distance increases. However, one sample t tests between each actual distance ratio and each imaged distance ratio showed that participants imaged distance ratios were significantly smaller than actual distance ratios ($p$s < .001, see Table 2.2 for details), which demonstrates that participants significantly underestimated the distance between landmarks when shifting attention across a mental image in this task. This was indicated for all age groups at all ratios, apart from 6- to 7-year-olds and 8- to 9-year-olds who showed no difference between the first actual distance ratio and first imaged distance ratio and 8-9 years additionally showed no difference between second actual distance ratio and second imaged distance ratio (*p*s > .05).

Table 2.2

*One sample t tests between actual distance and imaged distance ratio*

| *Imaged distances* | 6-7 years | 8-9 years | 10-11 years | Adults |
|---|---|---|---|---|
| | *t*(df), *p*, cohens *d* | | | |
| Difference from actual ratio 1.16 | -0.82(30), ns, 0.15 | -0.96(25), ns, 0.19 | -4.57(34), **, 0.77 | -3.69(57), ***, 0.48 |
| Difference from actual ratio 1.43 | -4.91(30), ***, 0.88 | -1.91(25), ns, 0.37 | -8.36(34), ***, 1.46 | -9.88(57), ***, 1.29 |
| Difference from actual ratio 2.2 | -11.93(30), ***, 2.14 | -11.58(25), ***, 2.27 | -13.12(34), ***, 2.22 | -13.43(57), ***, 1.76 |
| Difference from actual ratio 3.74 | -10.21(30), ***, 1.83 | -12.07(25), ***, 2.37 | -19.73(34), ***, 3.33 | -16.09(57), ***, 2.11 |

To examine whether deviation from actual distance is greater as the distance increases, difference scores were calculated between imaged and actual distance ratios and a mixed ANOVA was conducted to test whether participants' imaged distances deviated significantly further from actual distance at greater distances, and whether this effect was present in all age groups. There was no significant main effect of age ($F < 2$, $p > .05$). There was a significant main effect of distance ($F(1.77,257.67) = 530.63$, $p < .001$, $\eta_p^2 = .78$) which is best explained within the context of the interaction of distance and age group,($F(5.29, 257.66) = 4.09$, $p < .01$, $\eta_p^2 = .08$). Follow up ANOVAs revealed significant main effects of distance for all age groups (6- to 7-year-olds: $F(1.58,47.29) = 77.19$, $p < .001$, $\eta_p^2 = .71$; 8- to 9-year-olds: $F(1.89,47.27) = 108.47$, $p < .001$, $\eta_p^2 = .81$; 10- to 11-year-olds: $F(2.03, 68.96) = 249.51$, $p < .001$, $\eta_p^2 = .88$; Adults: $F(1.71, 97.48) = 191.90$, $p < .001$, $\eta_p^2 = .77$). With the exception of the 8-9-year-olds, all pairwise comparisons were significant (*p*s < .05 for all) indicating that as distance increased, deviation from distance increased. For children age 8-9 years, all comparisons were significant ($p < .05$ for all) except for between the first difference score and second difference score ($p > .05$). The equivalent analyses were conducted on perception control trials and the same relationships were found (see **Appendix** A.1.2.).

### 2.3.6. The relationship between each component of MI, VWM and attention control

For the correlational analyses reported below, it was important to use a variable that was sensitive to participants' ability to represent distance in order to fully capture image scanning ability. Therefore, for the purpose of this analysis an index of metric properties (IMP) score was calculated whereby the mean of each imaged distance ratio was subtracted from each actual distance ratio and squared, thus the closer the score to 0, the lower the deviation from actual distance. The following variables were included in the correlation analyses: overall accuracy in image generation, overall accuracy in image maintenance, mental rotation mean RT correct trials, image scanning mean IMP, VWM maintenance (forward VWM span), VWM manipulation (backward VWM span) and Go/No-Go mean correct RT. Descriptives statistics of all variables included in the correlation analyses are reported in Table 2.3. For reference, age group comparisons of VWM maintenance, VWM manipulation and Go/No-Go correct RT (attention control) are reported in the **Appendix, A.2.4.** As most variables were not normally distributed, Spearman's correlations were conducted. For the child data, a Spearman correlation matrix on the residuals of the ranks of the variables (age partialled out) was conducted. A separate Spearman's correlation matrix was conducted with the adult group.

Table 2.3

*Descriptive statistics for each of the variables included in the correlation analyses per age group*

| Measures | 6-7 years | 8-9 years | 10-11 years | Adults |
|---|---|---|---|---|
| | | Mean (SD) | | |
| Image generation overall accuracy | 13.00 (3.26) | 14.19 (3.20) | 15.97 (3.19) | 15.98 (3.55) |
| Image maintenance overall accuracy | 14.36 (3.61) | 14.89 (3.08) | 16.54 (2.67) | 17.97 (2.56) |
| Mental rotation correct RT (secs) | 4.77 (1.24) | 4.67 (1.72) | 3.35 (1.12) | 2.13 (.65) |
| Image scanning IMP | 1.56 (.76) | 1.50 (.92) | 1.39 (.72) | 1.06 (.54) |
| VWM maintenance (forward span) | 4.07 (.77) | 4.15 (.83) | 4.74 (.89) | 5.67 (1.20) |
| VWM manipulation (backward span) | 3.61 (.88) | 3.96 (.99) | 4.69 (.83) | 5.56 (.89) |
| Go/No-Go correct RT (secs) | .69 (.11) | .61 (.09) | .56 (.09) | .44 (07) |

As outlined in Table 2.4, after correction for multiple comparisons, in children aged 6-11 years, there were no significant associations between any of the components of MI, suggesting support for a separable component model (*p*s > .002, multiple comparison cut-off). Additionally, there were no relationships between either VWM maintenance or VWM manipulation and each of the components (*p*s > .002). With regards to attention control and components of MI, a positive relationship between Go/No-Go correct RT and mental rotation correct RT was found ($r_s$ = .378, $p$ < .001). The same correlation matrix was conducted with slope values as the mental rotation and image scanning variables. Findings indicated no significant associations between each of the variables (*ps* > .002) including between Go-No/Go correct RT and mental rotation slopes ($r_s$ = .131, $p$ = .217). For reference, the correlation matrix was also conducted with high precision mean RT for image generation and for image maintenance. All relationships remained the same in that there were no significant associations between high precision mean RT for image generation and image maintenance and the other components of MI or the components of VWM. However, there was a significant positive relationship between high precision mean RT in image generation and image maintenance ($r_s$ = .495, $p$ < .001).

Table 2.4

*Spearman's correlations on the residuals on the ranks of the variables (age partialled out) between components of MI, VWM and attention control in children age 6-11 years (N=92)*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. |
|---|---|---|---|---|---|---|---|
| 1. Image generation | — | | | | | | |
| 2. Image maintenance | .242 | — | | | | | |
| 3. Mental rotation | -.043 | -.200 | — | | | | |
| 4. Image scanning | -.052 | .092 | -.029 | — | | | |
| 5. VWM maintenance | -.036 | .112 | -.049 | .107 | — | | |
| 6. VWM manipulation | .128 | .009 | -.136 | -.137 | .225 | — | |
| 7. Go/No-Go | .066 | -.141 | .378*** | -.070 | -.074 | -.104 | — |

*** $p < .002$

How abilities in components of MI are related in adulthood was then examined. These findings were different to those found in primary school children. After correcting for multiple comparisons, there was a negative correlation between image maintenance and mental rotation RT ($r_s$ = -.487, $p$ < .001), suggesting higher accuracy in image maintenance is associated with faster correct RT in mental rotation. There were no relationships between measures of MI and VWM in adults. Finally, there was a positive association between the maintenance and manipulation measures of VWM in adulthood ($r_s$= .463, $p$ < .001). Despite moderate correlations between attention control and image generation (*rs* = -.358, $p$ = .007), image maintenance (*rs* = -.361, $p$ = .006) and mental rotation (*rs* = .372, $p$ = .005), these correlations did not survive corrections for multiple comparisons ($p$ < .002) (Table 2.4). This correlation matrix was conducted again with the mental rotation and image scanning slopes values. In this analysis, the relationship between image maintenance and mental rotation was more moderate ($r_s$ = .299, $p$ = .025). All other relationships remained the same. As in the correlation matrices with child data, the adult correlation matrix was conducted with high precision mean RT in image generation and image maintenance, neither variable was significantly associated with any of the components of MI or VWM (*p*s > .002).

Table 2.5

*Spearman's correlations between components of MI, VWM and attention control in adults (N=57)*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. |
|---|---|---|---|---|---|---|---|
| 1. Image generation | — | | | | | | |
| 2. Image maintenance | .346 | — | | | | | |
| 3. Mental rotation | -.219 | -.487*** | — | | | | |
| 4. Image scanning | -.154 | -.127 | .002 | — | | | |
| 5. VWM maintenance | .007 | .301 | -.228 | .020 | — | | |
| 6. VWM manipulation | .264 | .275 | -.299 | -.093 | .463*** | — | |
| 7. Go-No/Go | -.358 | -.361 | .372 | .045 | .028 | -.101 | — |

*** $p < .002$

## 2.4. Discussion

The contributions of this chapter are threefold. The first aim was to derive a novel battery of MI tasks suitable for capturing individual differences in visually depictive mental images in childhood and adulthood. The second aim was to examine evidence for a sub-component model of MI in development and the third aim was to investigate how components of MI relate to VWM and attention control abilities in childhood and adulthood. With reference to the first aim, it was demonstrated that participants of all ages tested were able to generate and maintain images of high visual precision. The image scanning task also provided novel evidence regarding the ability to represent distance in visual mental images in both children and adults in that as distance scanned increased, the extent to which participants underestimate distance was greater. With reference to the second aim, contrary to expectations, evidence for a separable-component model of MI was supported in childhood in that there were no significant associations between components of MI. Components appear to become more integrated in adulthood; demonstrated by a significant negative association between image maintenance accuracy and mental rotation RT in the adult group. With reference to the third aim, the evidence suggests MI, VWM and attention control were broadly dissociable in children and adults. The evidence for each of the aims will be discussed in turn in the following sections.

### 2.4.1. Evidence for visually depictive mental images in childhood

The findings presented in this chapter extend previous research by establishing the extent to which a depictive theory of MI is supported in childhood; specifically, it was found visual images are generated and maintained with precision throughout childhood and in adulthood. Participants of all ages indicated faster responses in high precision trials compared to low precision trials in both image generation and image maintenance. This supports the hypothesis that in the instance of generation and/or maintenance of low precision images, individuals are likely searching for an alternate, less efficient strategy, such as verbal strategies. With regard to developmental progression, adult levels of ability were reached at around 8-9 years in the image generation task. In the image maintenance task, the ability to maintain an image of high precision was found to be developing up to later primary school years (age 10-11 years). This is later than shown in Wimmer et al. (2015). Moreover, in contrast to Wimmer et al. (2015), higher accuracy and faster RT in image maintenance compared to image generation was only found in the adult group. This is likely dependent on the fact that the image maintenance task presented here requires a visual image of high precision to score highly. The use of abstract shapes in the current image generation and image maintenance tasks means that it is more difficult to score highly if individuals are relying on less visually precise mental images or strategies other than MI, such as verbalisation, as the stimuli are not easy to label verbally. Thus, it may be that the ability to maintain a highly precise visual image is still developing into later childhood, whereas the ability to remember locations of known objects or visuo-spatial imagery, as in Wimmer et al. (2015), requires a less visually precise image and reaches adult-like levels earlier in childhood.

Support for this suggestion comes from research investigating development of precision in VWM maintenance. In a study involving 90 males aged 7-13 years, researchers presented an orientated, coloured bar, which participants were required to memorise, hold in mind (500ms) and subsequently rotate a dial to move a probe stimulus to the position of the remembered stimulus. Varying degrees of distance between the remembered and probe stimulus allowed for measurement of precision of the maintained stimulus. It was found that precision in VWM maintenance continues to develop between age 7 and 13 years (Burnett Heyes et al., 2012). Overall, the findings presented here, which indicate later development of image maintenance in this task that requires visual images of high precision, is in line with the suggestion that the precision at which visual information is held in mind is developing throughout childhood beyond the age range tested in the current study.

Evidence of a linear time-degree of rotation effect in mental rotation and a linear time-distance effect in image scanning in all age groups is in line with previous findings that children from age 4 demonstrate evidence for image transformation ability and for the ability to shifting attention across an image (Estes, 1998; Frick et al., 2014; Möhring et al., 2016; Wimmer et al., 2016, 2017). Analysis of the slope values in mental rotation showed a steeper slope in the 8-9-year-old group compared to adults, again it is currently unclear whether this effect might be attributed to the particular sample of 8-9-year-olds. In line with Wimmer et al. (2016), there was no effect of age on steepness of slope on image scanning, suggesting that the ability to shift attention across an image is present from early childhood.

Current knowledge regarding individual differences in the precision of visual mental images in an image scanning task is extended in this study by demonstrating how accurately individuals mentally represent distances in the absence of sensory stimuli. This analysis revealed that when engaging in image scanning, participants underestimated actual distances, and that this deviation in imaged distance from the actual distance ratios increased with greater scanning distances. Previous research has evaluated perceived distances relative to the self. In an eyes-closed condition, it was found that individuals underestimate further distances compared to nearer distances (Fukusima et al., 1997). In eyes-open distance perception, a recent study found that 50% of participants consistently underestimated across all distance ratios (Norman et al., 2016). Moreover, distance estimation studied in children tends to show that primary aged children are less accurate than adults (Giovannini et al., 2009; Thurley & Schild, 2018). Considered alongside the evidence presented in the current study showing that individuals' perceived distance ratios mirror individuals' imaged distances, it appears that individuals make similar errors in internally representing distance in the absence of sensory stimuli as they do in distance perception. However, it should be noted that imagining distance and perceiving distance are not entirely aligned in that imaged ratios were lower than perceived ratios, as reported in the **Appendix, A.1.3**. In addition, inspection of the descriptive statistics for the index of metric properties scores shows that the minimum score was .06, whereas the maximum score was 3.24. Here, a score of 0 would indicate imaged distance ratios aligned exactly with actual distance ratios. Therefore, some individuals demonstrate highly precise visual images in the image scanning task, in line with evidence for percept-like, visually depictive representations in MI (e.g. Dijkstra et al., 2019; Ganis, 2013; Kosslyn et al., 2006). Yet, it is important to note that the range above indicates large individual differences in the ability to represent distances accurately.

### 2.4.2. Evidence for a separable-component model of MI

Evidence for a separable-component model of MI was supported, except some indication of a relationship between image generation and image maintenance. There was some indication of a moderate positive correlation between image generation and image maintenance in children, which was not significant between the accuracy measures but was significant between the RT measures. On the other hand, in adults there was no significant correlation between the accuracy measures or RT measures of image generation and image maintenance. These findings could be dependent on developmental differences in speed of processing between children and adults (e.g., Kail, 1991). Further investigation into the approach to image generation vs. image maintenance tasks is required to support this conclusion. It might be argued that the correlation between image generation and image maintenance is masked in the childhood sample because the image generation task requires switching to a distractor task in the trial, which might involve other developing cognitive processes such as switching. However, the data do not support this assertion because there were no significant differences in high precision responses between the two tasks in children (if switching had impacted performance, performance would have been poorer for image generation than image maintenance). Moreover, it is important to note that while separate correlation analyses per age group were planned based on medium effect sizes in previous research (Wimmer et al., 2015), sensitivity power analysis (detailed in **Appendix A.1.1.**) showed the individual age groups were not powered to detect weak correlations. Thus, while this evidence supports evidence for separable-component model of MI, it is not possible to delineate how the components of MI are related to one another throughout the primary school years based on this data. Given the previous conflicting findings regarding support for a separable-component model in different age groups (Kosslyn et al., 1990; Wimmer et al., 2015), future research powered to detect weak correlations is warranted.

Furthermore, in the adult group, image maintenance accuracy and mental rotation RT are significantly negatively correlated suggesting that as accuracy in image maintenance increases, RT in correct trials of mental rotation decreases. This suggests that holding an image in mind is important for both sub-components of MI. However, it should be noted that the relationship between mental rotation slope values and image maintenance accuracy was more moderate and fell short of the multiple comparison cut-off. It could be argued that the slope values are more sensitive to transformation of representations, whereas correct trial RT is more sensitive to holding a visual representation in mind, hence the significant association with image maintenance. Firstly, it is important to note that conclusions should be tentative given the inflation of type II error with Bonferroni corrections (Perneger, 1998). That being said, a lack of association between mental rotation slope values and image maintenance accuracy appears to be supported by a similar dissociation in individuals with Aphantasia (i.e., individuals without the ability to generate visual images; Zeman et al., 2015). These individuals present with intact subjectively reported spatial imagery, i.e., transforming representations, alongside absent visual imagery (Dawes et al., 2020; Keogh & Pearson, 2017) and one study has shown mental rotation performance is on par with controls (Pounder et al., 2018). On the other hand, there is also evidence to suggest visual representations are recruited and held in mind during mental rotation (Hyun & Luck, 2007; Prime & Jolicoeur, 2010), which might explain why image maintenance accuracy was significantly associated with mental rotation correct RT.

Clearly more research is required; whilst the childhood data in the current study supports previous evidence of a dissociation between mental rotation and image maintenance (Keogh & Pearson, 2017a; Kosslyn et al., 1990), the adult data is not so clear cut. The suggestion that components of MI in adulthood are not entirely separable is not new. Specifically, the seminal paper investigating how activation in different areas of the brain predicted performance in different imagery tasks (image generation, image inspection, image transformation, image resolution) found that whilst there was evidence to suggest differential activation between imagery tasks, there was also some evidence for shared activation between tasks that require maintaining an image and tasks that require transforming an image in the occipito-parietal sulcus, medial frontal cortex and the early visual areas (Kosslyn et al., 2004). Thus, it is tentatively concluded, based on the association between image maintenance and mental rotation that visual imagery and transformation abilities are at least partially integrated in adults.

### 2.4.3. Evidence for dissociated components of MI, VWM and attention control

Despite expectations based on previous evidence with adults, the current study implies ability in each of the components of MI, VWM and attention control are not related. First it should be noted that the positive association between VWM maintenance and manipulation in adults suggests they are not entirely separable, this is in line with a recent review (Donolato et al., 2017). The evidence presented in the current study suggests the components of MI and maintenance and manipulation in VWM are dissociable in childhood and adulthood. This is surprising given the evidence in adult research demonstrating a relationship between MI and VWM, even leading some to argue that they are not distinguishable (Tong, 2013). While it should not be claimed that components of MI and VWM are wholly dissociated based on this data, the findings do lend support to the argument that there are individual differences in the types of strategies employed in VWM tasks.

Research determining children's strategies in VWM is limited. If strategies are investigated, the relationship tends to be examined in the opposite direction, i.e., how can different strategy conditions influence VWM and how does this transfer to other abilities, such as mathematics (e.g., Cragg et al., 2017; Swanson, 2015). Studies where children are asked to reflect on their strategies are less common. In a study investigating transfer of gains in working memory training, 37% of the sample of children age 8-11 years retrospectively reported concentrating harder and 27% of participants reported a range of other strategies, including rehearsing information and tracing the pattern on the screen with their eyes (Holmes et al., 2009). The variability in participants' responses supports the argument above that a lack of relationship between MI and VWM measures in children may be dependent on ongoing development of precision of MI, and thus children may be reverting to less efficient strategies to complete VWM maintenance and manipulation tasks. Thus, it may be that while tasks are designed to measure *visual* WM specifically, there may be variation at the individual level as to whether participants are recruiting MI strategies. Future research should examine types of strategies children recruit in VWM and should assess developmental differences in the relationships between components of MI and VWM between narrower age groups (e.g., between 6-7-year-olds and 8-9-year-olds) with a sufficient sample size for small effect sizes.

In the adult literature, a study introducing irrelevant visual information to try and disrupt performance in VWM found only individuals who scored highly on the MI strength measure, i.e., "good imagers", were disrupted (Keogh & Pearson, 2014), thus implying variance in strategy-use during VWM may be dependent on individual differences in MI ability. Moreover, it has been suggested in adult literature that understanding individual variability in strategies and the format of representations in VWM might explain discrepancies between studies investigating neural mechanisms of VWM and why the numbers of items held in VWM can vary greatly (Pearson & Keogh, 2019; Reeder, 2017). The previous research with adults, and the research presented in this chapter, have made interpretations regarding strategy based on evidence comparing performance on MI tasks and VWM tasks. While inferences can be made, it is not possible to determine how MI strategies influence VWM performance without explicitly testing the extent to which visual strategies are recruited within a VWM task. Therefore, to provide vital evidence of how MI is recruited *within* a VWM task, **Chapter 4** will directly examine how individual differences in the vividness of mental images and the number of items held in mind impacts behavioural and neural correlates of VWM. This will build on findings in the current chapter comparing MI and VWM to address how MI is recruited *within* a VWM task.

With regards to attention control, different relationships were indicated in primary school children compared to adults. It should first be noted that correlations between attention control and the MI and VWM measures did not survive correction for multiple comparisons in either the child or adult correlation matrices. However, in the adult group, there were moderate correlations between attention control and image generation, image maintenance and mental rotation, which were just above the multiple comparison cut-off of $p$ = .002 ($p$s = .005 to .007). It has been argued that the Bonferoni corrections can be too conservative and results in higher likelihood of type II errors (Perneger, 1998). Considering this alongside the evidence that the effect sizes were moderate (Cohen, 1988), the correlations between MI measures and attention control in the adult group are interpreted below with caution. Firstly, in the primary school sample, the attention control measure derived from Go/No-Go correct RT was significantly and positively related to mental rotation correct RT. However, given that there was no relationship between attention control and mental rotation slope values, the significant correlation between the two correct RT measures likely reflects a relationship between processing speed in the two tasks. Interestingly, despite almost non-existent correlations (very low r values) between the attention control measure and components of MI in the primary school sample, the adult sample appeared to show the opposite effect. Namely, there were moderate correlations between attention control and image generation, image maintenance and mental rotation, which were just above the multiple comparison cut-off ($p$s = .005 to .007). Firstly, the finding that there was no relationship between image scanning and attention control should be addressed. This might seem counter-intuitive at first glance given that image scanning is defined as the ability to shift attention across a mental image. However, this is not a direct measure of attention control, and the image scanning variable in the current study was specifically derived to measure the ability to represent varying distances in a visually depictive mental image, therefore it is perhaps not surprising that image scanning is not associated with attention control.

The moderate correlations denoted in the adult group can be tentatively interpreted given the size of the coefficient. Broadly, it would appear that attention and MI abilities become integrated in adulthood following development as distinct abilities. Firstly, the finding that MI and attention control are associated in adults supports previous literature suggesting that vivid colour imagery can prime subsequent attention selection (Cochrane et al., 2020). While consideration of the role of attention control in MI is limited, the theoretical perspective of a focus of attention within working memory has been comprehensively formulated (Cowan, 2001; Cowan, 2011, 2014, 2016; Cowan et al., 2005). With regard to development of working memory, it is argued that children's poorer ability to maintain visual information in mind is dependent on a limited ability to focus attention rather than limited capacity. Therefore, greater performance in VWM observed in older participants is likely due to greater attention control and thus more stable representations (Cowan, 2016; Shimi et al., 2014). A similar attention-processing mechanism might be at play in image generation, maintenance, and mental rotation of visually depictive mental images. In line with the attention-processing account of working memory, the correlational data presented in this study imply that the more visually precise the image is (as characterised by more precise images in the adult group compared to the youngest age group), the greater the attention control required. Nevertheless, it is important to note here that a more sensitive measure of attention control would be number of errors, however because performance was at ceiling, the measure was derived from correct RT. Moreover, findings regarding the development of a focus of attention in working memory have shown children above age 7 show adult-like abilities, therefore further research with larger sample sizes per age group to detect moderate coefficients is required to examine how attention control is related to MI components at different stages of development.

### 2.4.4. Conclusion

In conclusion, this chapter has demonstrated that children from at least age 6 are able to generate and maintain highly precise visual images of abstract shapes, thus providing support for a depictive theory of MI in childhood. Linear time-angle effects and time-distance effects in mental rotation and image scanning from age 6 are also replicated in this study. Novel insights into image scanning abilities are provided by evaluating how individuals internally represent distance in the image scanning task in both children and adults, which demonstrated, as in distance perception, individuals underestimate distance as distance increases. Importantly, the tasks presented in this study extend previous methodological restrictions by firstly providing a battery of tasks suitable for quantifying MI abilities in both primary school aged children and adults, and secondly, allowing for the evaluation of individual differences in visually depictive mental images. Support for a separable-component model of MI in development was demonstrated alongside evidence for integration between some components in adulthood. It should be noted that while we can make inferences about development, this initial study involves a cross-sectional sample and longitudinal research is required to provide clarity to developmental findings.

Nevertheless, this study is a valuable starting point in understanding the format of representations in children's thinking and should lead to research on how individual differences in representational formats, in both MI and VWM respectively, develop over primary years and how this might impact learning. A vital step in addressing how within-task variability in visually depictive mental images contributes to VWM ability is outlined in **Chapter 4.** As the first study to examine the role of attention control in components of MI, it can be tentatively concluded that an attention-processing mechanism might be at play in the ability to generate, maintain and transform visually depictive mental images in adults. However, it should be recognised that the correlations in the adult group did not survive corrections for multiple comparisons. This has important implications for how MI abilities present in children with attention deficits, which is directly addressed in **Chapter 3.** The findings also bring to light the value of investigating separable components of MI, VWM and attention in development and provide a crucial contribution to the theoretical perspective that MI develops as a multi-faceted function.

# Chapter 3: Characterising MI abilities alongside VWM in children with ADHD: examining group-level and individual-level effects

## 3.1. Introduction

A principal aim of this thesis is to characterise MI abilities in children with ADHD. As outlined in **Chapter 1**, on average, children with ADHD present with impairments in VWM compared to their typical peers. VWM is positively also associated with greater academic achievement in typical development and positively associated with poorer academic outcomes in ADHD. Despite this working memory training appears to be ineffective in both TD children and children with ADHD. Moreover, more recent research suggests there are in fact extensive individual differences in VWM abilities in both typically developing (TD) children and children with ADHD. In order to support children in their learning, it is vital to understand the mechanisms underpinning abilities in VWM. In parallel literature with adult populations, the relationship between MI and VWM has been investigated and recent reviews have argued individual variability in the recruitment of MI strategies might explain variability in capacity limits in VWM. Thus, to establish a potential source of individual differences in VWM in children with ADHD, it is vital to address these gaps in the literature and examine (1) how MI presents in children with ADHD and (2) how MI abilities relate to VWM abilities in children with ADHD and TD children. These form the aims of this chapter.

### 3.1.1. ADHD and working memory

Research reviewed in **Chapter 1, Section 1.3.1.** indicates that children with ADHD have poorer working memory than their TD peers, with greater deficits in the visual domain compared to the verbal domain (Gau et al., 2009; Gau & Chiang, 2013; Martinussen et al., 2005; Martinussen & Tannock, 2006; Nikolas & Nigg, 2013; Simone et al., 2015). This is specifically demonstrated in separable tasks requiring maintenance and manipulation of visual information in mind (Martinussen et al. 2005; Simone et al., 2015; Kasper et al., 2012). This has led some to argue that deficits in working memory are a defining neuropsychological attribute of ADHD (Alderson et al., 2010), whereas more recent research has indicated that poor working memory is largely moderated by factors other than ADHD severity, such as learning difficulties (Nikolas & Nigg, 2015). Moreover, data-driven analysis techniques have revealed that there is extensive variability in VWM abilities in both children with ADHD and TD children (Campez et al., 2020; Fair et al 2011; Kofler et al., 2019). While there is strong evidence that VWM is positively associated with academic outcomes in both TD children and children with ADHD (Best et al., 2011; Friedman et al., 2018; Orban et al., 2018; Simone et al., 2018), the relationship between ADHD severity, working memory and academic outcomes is unclear. Evidence suggests that ADHD symptoms significantly predict working memory performance, however this is not evident for all measures of VWM (Martinussen et al., 2005; Tillman et al., 2011). With regards to ADHD symptoms and academic achievement, a recent study involving a range of academic outcome measures, including reading comprehension, mathematics and spelling, found that VWM contributed to each of the outcomes, but that the severity of ADHD symptoms did not moderate these relationships (Simone et al., 2018). Moreover, despite positive associations, training VWM in children with ADHD and TD children has little effect on academic outcomes (Cortese et al., 2015; Rapport et al., 2013; Sala & Gobet, 2017). Taken together, it is not possible based on the current evidence to conclude that impaired VWM is a defining attribute of ADHD nor is it possible to determine how VWM contributes to academic outcomes in ADHD. Prior to establishing why VWM training is ineffective, it is vital to take a step back and examine the role of individual differences in VWM in ADHD.

### 3.1.2. ADHD, MI and VWM

Conceptualisations of MI and VWM tend to overlap; both involve the maintenance and manipulation of visual information (Baddeley, 2003; Baddeley & Andrade, 2000; Cowan, 2001; Logie, 1995). This has led some to argue that the two are in fact synonymous (Tong, 2013). However, the current evidence supports a more moderate argument that while there are shared mechanisms between MI and VWM, they are not one and the same. Firstly, there is evidence for shared representations in the early visual areas between MI and VWM, in that it is possible to decode neural activity in the early visual areas during MI using a classifier trained on VWM trials and vice versa (Albers et al., 2013). This suggests that both processes recruit visual representations. However, the effective use of such visual representations may depend on individual differences. Namely, it has been suggested that only those with strong imagery (i.e., scoring highly on a sensory strength measure of MI) appear to recruit visual strategies in VWM, whereas those with weaker imagery might rely on symbolic/language-like representations or verbal strategies akin to general thought (Keogh & Pearson, 2011; 2014; Pearson & Keogh, 2019). Moreover, **Chapter 2** presents the first known study to investigate how components of MI and VWM are related to one another, and it was found that MI and VWM appear to be dissociable in primary school children and in adults. This importantly highlights that the relationship between MI and VWM is complex. Despite evidence for heterogeneity in VWM abilities in children with ADHD, research into how MI abilities present in children with ADHD is limited. Thus far, only mental rotation has been examined in children with ADHD, therefore the picture of MI abilities in ADHD is incomplete. The few studies that have investigated how children with ADHD perform on mental rotation tasks have found impaired mental rotation ability in children with ADHD compared to TD controls (Jakobson & Kikas, 2007; Silk et al., 2005; Vance et al., 2007; J. Williams et al., 2013). However, it is important to note that these studies tend to include very small sample sizes: N range = 7 to 26. Therefore, further research is required to examine the full range of mental rotation abilities in a representative sample of children with ADHD. Overall, the limited investigation into MI abilities in children with ADHD presents a clear gap in the literature aiming to determine the neuropsychological profile of ADHD.

### 3.1.3. The current study

The current study was designed to characterise MI abilities alongside VWM in children with ADHD and TD children and to assess the relationship between MI and VWM in these groups. Firstly, a case-control design was adopted to compare abilities in both groups. As outlined above and reviewed in detail in **Chapter 1, Section 1.3.**, only examining between-group differences limits understanding of cognitive functions in ADHD and masks important individual differences in both TD children and children with ADHD. Recent research has therefore adopted transdiagnostic approaches to augment our understanding of individual differences beyond what can be gleaned from comparing group means (see Astle et al., 2021 for review). With regard to VWM, data-driven, transdiagnostic approaches have provided evidence for a wide range of VWM abilities in both TD children and children with ADHD (Campez et al., 2020; Kofler et al., 2019; Dajani et al., 2014). Data-driven clustering techniques, such as latent profile analysis, allow us to identify profiles of cognitive ability derived from individual differences rather than diagnostic labels. The use of this technique in the current study alongside between-group comparisons has theoretical importance. While previous evidence with adults has suggested a positive relationship between MI and VWM abilities in adults, the findings presented in this thesis thus far suggest a dissociation between MI and VWM in both children and adults. Therefore, data-driven analysis will not only provide understanding on the range of abilities in both TD children and children with ADHD but will examine how individual differences contribute to profiles of MI and VWM abilities. Therefore, a latent profile analysis was conducted to investigate individual differences in MI and VWM in both TD children and children with ADHD to derive distinct profiles of abilities beyond diagnostic labels.

Tasks measuring each of the MI components introduced in **Chapter 2** were adopted in this study. Given that between-group comparisons tend to indicate poorer performance in VWM in ADHD groups compared to TD groups and there is evidence in the literature for shared mechanisms between MI and VWM, it is hypothesised that children with ADHD will present with poor MI relative to TD children of the same age in all four components of MI. The components of image generation, image maintenance and image scanning are yet to be directly investigated in children with ADHD. Poorer abilities are predicted relative to chronological age; however, it is expected that children with ADHD will exhibit a typical pattern of performance. This is in line with what is observed in VWM, i.e., both children with ADHD and TD children are found to show a pattern of worse performance in backward span tasks (VWM manipulation) compared to forward span tasks (VWM maintenance), however overall performance is poorer in children with ADHD compared to age-matched TD children (Martinussen et al., 2005; Martinussen & Tannock, 2006; Simone et al., 2015). Therefore, in line with findings in **Chapter 2**, it is predicted that children with ADHD will demonstrate faster RTs in high precision compared to low precision trials in image generation and image maintenance, respectively. This would suggest that children with ADHD present a typical pattern of results in image generation and image maintenance, however their performance is expected to be poor for their age as seen in VWM research. In the image scanning task, the dependent variables are derived from RT therefore poorer performance compared to chronological age is expected, as well as an atypical pattern of performance. This is expected in light of evidence suggesting that greater variability in RTs compared to typical peers is a common characteristic of children with ADHD (Kofler et al., 2013).

With regard to mental rotation, previous research has shown children with ADHD demonstrate significantly poorer performance compared to chronological age-matched controls (Silk et al., 2005; Vance et al., 2007; Williams et al., 2013; Jakobson & Kikas, 2007) and differential patterns of performance with regards to RT measures in mental rotation have also been indicated (Feldman & Huang-Pollock, 2020). Therefore, it is predicted that in the current study, children with ADHD will exhibit low accuracy for their age, and atypical patterns of performance in RT measures in the mental rotation task. In line with previous evidence for impairments in VWM in children with ADHD, it is predicted that group-level analyses will demonstrate poorer VWM in children with ADHD compared to TD children of the same age. However, given the evidence for individual variation in VWM abilities in children with ADHD and TD children, individual differences are expected in the latent profile analysis for all children. As individual differences in VWM are yet to be investigated alongside individual differences in MI in either TD children or children with ADHD, there are no predefined hypotheses as to what profiles of ability will arise in the latent profile analysis involving all participants.

The second aim of this study is to examine how components of MI relate to one another, to maintenance and manipulation measures of VWM and to symptoms of ADHD in children with ADHD compared to TD children. The VWM tasks, forward span (VWM maintenance) and backward span (VWM manipulation) were the same as those employed in **Chapter 2.** The findings from the primary school sample in **Chapter 2** indicated that that there were no associations between each of the components of MI, supporting a separable component model of MI, and there were no associations between each of the components of MI and either of the VWM measures. In light of heterogeneity of VWM in children with ADHD and the lack of previous research examining MI in ADHD, there are no predefined hypotheses as to how or whether components of MI will be related to maintenance and manipulation in VWM. Finally, while the relationship between VWM and ADHD symptoms has been previously examined (Martinussen et al., 2005; Tillman et al., 2011), research is yet to investigate how ADHD symptoms are associated with components of MI in either TD children or children with ADHD. It is vitally important to understand how neuropsychological profiles might interact with behavioural presentation of ADHD symptoms, therefore this study will clarify how ADHD symptoms are associated with components of MI and measures of VWM maintenance and manipulation.

## 3.2. Methods

### 3.2.2. Participants

Twenty-one participants with ADHD were recruited online via social media, in person via support groups for parents with children with ADHD and through primary schools in London and the surrounding areas. For children to be included in the ADHD group, parents were required to provide confirmation of diagnosis from their child's clinician, T scores on the Global ADHD Index derived from the Conners 3 Parent Rating Scale Long Form (CPRS-3:L) (Conners, 2008) were required to be ≥ 65 and if the child was taking medication for their ADHD symptoms, they were required to refrain from that medication for at least 24 hours prior to testing. In three instances, the Conners 3 Global ADHD Index T score was < 65, even though the parent provided confirmation of an ADHD diagnosis. In these cases, the Inattentive scale and Hyperactivity scale were scored, and the T scores were > 65, therefore these three participants were included in the ADHD group. Co-occurring disorders are common in ADHD, specifically autism spectrum disorder (ASD) which is diagnosed in 37-85% of children with ADHD (Leitner, 2014). For the findings to be generalisable to a representative population of children with ADHD, we did not exclude children with co-occurring disorders (ADHD only N = 11, ADHD and ASD N = 4, ADHD and anxiety, attachment disorder and dyslexia N = 1, ADHD and dyspraxia N = 1, ADHD and dyslexia N = 1). One participant was excluded as they did not complete the full battery of tasks and one more participant was excluded due to an incomplete parent consent form; the final ADHD sample consisted of 19 children. Data from the primary school sample in **Chapter 2** was used to derive the TD sample in the present study. To ensure this sample were typically developing, children who scored ≥ 65 on the T score from either the Inattentive or Hyperactivity/Impulsivity scales from the Conners 3 Teacher Scale Short Form (CTRS-3:S) (Conners, 2008) were excluded, which led to the exclusion of 23 of the sample of the 92 primary school children from **Chapter 2**. While it would have been preferable to use the same measure for ADHD symptoms in the TD sample, i.e., the Conners Global ADHD Index, it is not possible to derive from the CTRS-3:S. A total of 88 children were included in the final sample (6- to 7-year-olds: N = 27, female = 19; 8- to 9-year-olds: N = 19, female = 12; 10- to 11-year-olds: N = 23, female = 8; ADHD group: N = 19, female = 1). Participant demographics are reported in Table 3.1.

Table 3.1

*Demographics of TD and ADHD groups*

| Group | 6-7 years | 8-9 years | 10-11 years | ADHD |
|---|---|---|---|---|
| | Mean (SD) | | | |
| Age in years; months | 6;11 (0;06) | 8;04 (0;06) | 10;07 (0;03) | 11;00 (1;07) |
| RCPM Raw Score | 24.54 (4.55) | 26.71 (4.70) | 31.26 (2.59) | 32.33 (2.45) |
| CPRS-3:L Global ADHD Index T score | N/A | N/A | N/A | 77.90 (12.85) |
| CPRS-3:L Hyperactivity scale T score | N/A | N/A | N/A | 79.74 (14.17) |
| CPRS-3:L Inattentive scale T score | N/A | N/A | N/A | 81.05 (9.58) |
| CTRS-3:S Hyperactivity scale T score | 48.89 (10.29) | 47.37 (8.30) | 47.91 (8.43) | N/A |
| CTRS-3:S Inattentive scale T score | 50.93 (6.78) | 47.89 (8.04) | 46.96 (6.21) | N/A |

*Note.* RCPM = Raven's Coloured Progressive Matrices. Mean and SD RCPM represent twenty-six 6- to 7-year-olds and seventeen 8- to 9-year-olds on account of one missing RCPM in the 6- to 7-year-old group and two missing RCPM in the 8- to 9-year-old group, respectively. As CPRS-3:L was collected for the ADHD group and CTRS-3:S was collected for the TD group, N/As are included.

One-way ANOVAs were conducted on the background measures to compare groups. A significant main effect of age in years (*Welch's* $F(3,39.98) = 303.94$, p < .001, $\eta_p^2 = .80$) confirmed that the 10-11-year-old TD group were not significantly different in age from the ADHD group (post hoc comparison: $p = .584$), all other group comparisons were significant (*ps* < .001). Secondly, a one-way ANOVA of RCPM scores revealed a significant main effect of group (*Welch's* $F(3,41.27) = 21.705$, $p < .001$, $\eta_p^2 = .45$). Post hoc comparisons showed the ADHD group did not differ from 10-11-year-olds ($p = .535$), indicating mental age in the ADHD group was commensurate with their chronological age. All other comparisons were significant (*p*s < .001) indicative of developmental progression, except for the 6- to 7-year-old vs. 8- to 9-year-old groups ($p$ = .450).

### 3.2.3. Materials and procedure

The image generation, image maintenance, mental rotation, image scanning, forward span VWM and backward span VWM tasks introduced in **Chapter 2** were adopted in this study. TD children were tested at school and children with ADHD were tested either at the Institute of Education, University College London or at school in a quiet room. The order of tasks was counterbalanced across participants and instructions were administered exactly as in **Chapter 2**. The Conners 3 Rating Scales were additional to measures described in **Chapter 2** and are outlined in the section below. Finally, Raven's Coloured Progressive Matrices (RCPM) was carried out as a measure of non-verbal IQ (Raven et al., 1998).

#### 3.2.3.1. ADHD symptom measures

Conners 3 Rating Scales are widely used to measure symptoms of ADHD in children aged 6-18 years. As TD children were tested in school, teachers completed the CTRS-3:S. The short form was the most suitable option practically, given teacher's limited time, and the CTRS-3:S has been found to have good internal consistency (range for scales: .87 to .94) (Conners, 2008). The CTRS-3:S is made up of 40 items where teachers rate the child's behaviours over the last month on a scale from "not true at all" to "very much true (very often)". This provides 5 different symptoms scales: ADHD Inattentive, ADHD Hyperactive-Impulsive, ADHD Combined, Conduct Disorder and Oppositional Defiant Disorder. Raw scores were transformed into the standardised T scores based on age and gender for the ADHD Inattentive and ADHD Hyperactive-Impulsive symptom scales. As it is not possible to derive a Global ADHD Index score from the short form, the ADHD symptoms score for the correlation analyses was comprised of the child's highest score on either the Inattentive or Hyperactive-Impulsive symptom scale. The ADHD symptoms score was comprised of the highest score on either scale as opposed to a mean across the two scales because if the child scored predominantly Inattentive, e.g., T score = 80, then the Hyperactivity T score could be as low as 40, therefore a mean of the two scores would not necessarily flag a TD child with high ADHD symptoms. This is particularly important for the exclusion criteria. Descriptive statistics are reported in Table 3.1.

Parents of children with ADHD completed CPRS-3:L and were also given a Conners 3 Teacher Rating 3 Long Form (CTRS-3:L) to give to their child's teacher and then post back to the university. While it would have been preferred to have the same informant (teachers) for the TD and ADHD groups, only five out of eighteen CTRS-3:L were completed and mailed back, therefore for consistency, the CPRS-3:L was used for all ADHD participants. While low to moderate agreement between parent and teacher ratings has been indicated (A. L. Murray et al., 2018), research has also shown that there is no difference in diagnostic accuracy of ADHD between parent and teacher informants (Bied et al., 2017). The CPRS-3:L provides 14 subscales derived from a 108 item questionnaire with the same response structure as the CTRS-3:S. The Conners Global ADHD Index raw scores are calculated from 10 items and transformed into T scores based on age and gender where a score ≥ 65 is considered indicative of ADHD. The CTRS-3:L has good internal consistency (range of scales: .75-.94) and high test-retest reliability (range of scales: .71-.78) (Conners, 2008). The Conners Global ADHD Index T score was used as the ADHD symptoms dependent variable in the correlation analyses reported below and the descriptive statistics can be found in Table 3.1.

### 3.2.4. Analysis strategy

Analyses were preregistered (https://osf.io/7cyba/?view_only=2db4fe7c67cf417eb9d125d8deac08ed). As can be seen in the preregistration, the final sample size of children with ADHD is much lower than the planned sample size; this is due to fact that data collection was terminated 6 months early due to COVID-19 restrictions. The ANOVAs examining differences in response type between image generation and image maintenance, slope values for mental rotation and image scanning and mental rotation accuracy are additional to the preregistration. To retain as many participants in the ADHD group as possible, RCPM was not used as an additional exclusion criterion.

Tests of normality revealed most variables were not normally distributed, however parametric analyses were applied given that ANOVA is robust to violations of assumptions of normality (Blanca et al., 2017). Welch's *F* and Games-Howell corrected post-hoc comparisons are reported for one-way ANOVAs as there are unequal sample sizes across groups. Where estimates of sphericity were violated, Greenhouse-Geisser corrections are reported. All within-subject pairwise comparisons are reported with Bonferroni corrections. Levene's test of equality of variance of all RT variables exceeded appropriate assumptions for parametric tests in **Chapter 2**, however Levene's test was only violated in the mental rotation mean correct RT degree of rotation measures and the image scanning mean RT distance measures in this study. Therefore, raw RTs are reported in the image generation and image maintenance analyses alongside log-transformed RTs in the mental rotation and image scanning analyses. Levene's test was no longer significant in the log-transformed RT analyses in either mental rotation or image scanning (*p*s > .05). For the correlation analyses, Spearman's correlations are reported for the TD group due to nonnormal distributions of most variables (Field, 2013) and Kendal's tau was adopted for the ADHD group correlation analyses as it is more robust with smaller sample sizes (Gibbons, 1993). To correct for multiple comparisons, p values were Bonferroni corrected (alpha level 0.05 / 21 comparisons) resulting in *p* = .002.

## 3.3. Results

### 3.3.1. Image generation

To examine evidence for visually precise mental images in TD children and children with ADHD, the dependent variable of precision response type RT was derived as in **Chapter 2**. Precision response type has two levels: RT of high precision responses and RT of low precision responses. High precision responses are when the participant accurately identifies that the target shape is different and accurately identifies the exact location of the difference and is thus indicative of a highly precise visual mental image. Low precision responses are when the participant identifies that the target shape is different but does not correctly identify the exact location of the difference and is thus indicative of a visual mental image of low precision.

A mixed ANOVA was conducted on precision type RT (high precision, low precision) by group (TD 6-7-year-olds, TD 8-9-year-olds, TD 10-11-year-olds, ADHD group). There was no main effect of group ($F(3,84) = 1.94$, $p = .130$, $\eta_p^2 = .06$). There was a significant main effect of precision type RT ($F(1,84) = 37.03$, $p < .001$, $\eta_p^2 = .31$), which showed that participants were significantly faster to respond in high precision trials compared to low precision trials, see Figure 3.1. There was no significant interaction between response type RT and group ($F < 1$). This suggests children with ADHD are performing at an age-appropriate level and show a typical pattern of performance.

### 3.3.2. Image maintenance

A mixed ANOVA of precision response type RT (high, low) by group (TD 6-7-year-olds, TD 8-9-year-olds, TD 10-11-year-olds, ADHD group) was carried out, where precision response type was derived as in image generation. In contrast to the image generation analyses, there was a significant main effect of group ($F(3,84) = 6.43$, $p < .001$, $\eta_p^2 = .18$). This demonstrated that TD children aged 6-7 years were slower to respond than TD 8-9-year-olds ($p = .009$), TD 10-11-year-olds ($p < .002$) and the ADHD group ($p = .008$). All other comparisons were non-significant ($p > .05$ for all). There was also a significant main effect of precision response RT ($F(1,84) = 34.03$, $p < .001$, $\eta_p^2 = .28$) (Figure 3.1.), which revealed, in line with findings in image generation analyses above, significantly slower RTs in low precision responses compared to high precision responses. There was no significant interaction between precision response RT and group ($F(3,84) = 1.49$, $p = .222$, $\eta_p^2 = .05$). Therefore, children with ADHD show a typical pattern of ability in image maintenance, as well as performing at the level of TD children of the same age.

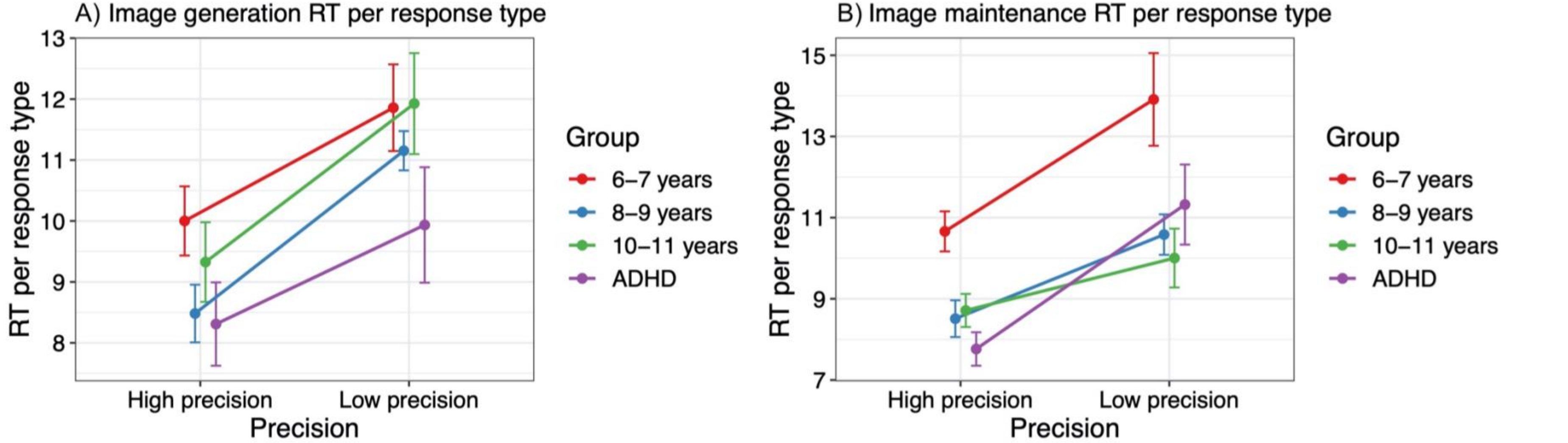



*Figure 3.1:* Mean and SE RT per level of precision (high, low) for image generation (A) and image maintenance (B)

### 3.3.3. Image generation vs. image maintenance

To examine how the relationship between image generation and image maintenance might differ between TD and ADHD groups, a mixed ANOVAs were conducted on task (image generation, image maintenance) by group for high precision trials only, where the dependent variable was RT. The ANOVA of response type RT showed a significant main effect of group ($F(3,84) = 5.49$, $p = .002$, $\eta_p^2 = .16$). Post hoc comparisons show 6- to 7-year-olds were significantly slower to respond than 8- to 9-year-olds ($p = .022$) and the ADHD group ($p = .002$) (all other comparisons, *p*s > .05). There was no main effect of task ($F < 1$) and no significant interaction ($F(3,84) = 1.02$, $p = .389$, $\eta_p^2 = .04$), thus showing that RTs did not differ between image generation and image maintenance tasks and children with ADHD presented the same pattern of results and at the same level as TD children of the same age.

### 3.3.4. Mental rotation

3.3.4.1. Mental rotation RT

A mixed ANOVA was conducted on the mean logRT of correct trials with a within-subject factor of degree of rotation (0°, 45°, 90°, 135° and 180°) and a between-subject factor of group. There was a significant main effect of group ($F(3,84) = 16.85$, $p < .001$, $\eta_p^2 = .37$) where post hoc comparisons revealed significantly slower responses in the TD 6- to 7-year-olds compared to the TD 10- to 11-year-olds ($p < .001$) and the ADHD group ($p < .001$), as well as between the TD 8- to 9-year-olds and the TD 10- to 11-year-olds ($p < .001$) and the ADHD group ($p < .001$). Crucially, the ADHD group did not differ from the TD 10- to 11-year-olds ($p = .969$). All other comparisons were non-significant ($p > .05$ for all). There was also a significant main effect of degree of rotation ($F(3.52,295.89) = 89.89$, $p < .001$, $\eta_p^2 = .52$) and a significant interaction between degree of rotation and group ($F(10.57,295.89) = 3.02$, $p = .001$, $\eta_p^2 = .09$). Follow up repeated measures (RM) ANOVAs revealed significant main effects of degree of rotation were best explained by significant linear contrasts in all groups (TD 6- to 7-year-olds: $F(1,26) = 21.79$, $p < .001$, $\eta_p^2 = .47$; TD 8- to 9-year-olds: $F(1,18) = 54.11$, $p < .001$, $\eta_p^2 = .75$; TD 10- to 11-year-olds: $F(1,22)$

= 121.59, $p < .001$, $\eta_p^2 = .85$; ADHD: $F(1,17) = 118.85$, $p < .001$, $\eta_p^2 = .87$) showing that as degree of rotation increased, RT increased. However, TD 6- to 7-year-olds also showed a significant quadratic contrast ($F(1,26) = 8.46$, $p = .007$, $\eta_p^2 = .25$) and the TD 8- to 9-year-old group revealed a significant cubic contrast ($F(1,18) = 6.21$, $p = .023$, $\eta_p^2 = 26$). Pairwise comparisons in the TD 6- to 7-year-old group revealed there was no significant increase in RT between incremental degrees of rotation (e.g., between 0° and 45°) (*p*s > .05) but there were significant increases between smaller degrees of rotation and larger degrees of rotation (0° < 90°, $p < .001$; 0° < 135°, $p < .001$; 0° < 180°, $p < .001$; 45° < 135°, $p < .001$) except for between 45° and 180° ($p = 1.00$). Pairwise comparisons in the TD 8- to 9-year-old group revealed significant increases between smaller degrees of rotation and larger degrees of rotation only (0° < 90°, $p = .014$; 0° < 135°, $p = .001$; 0° < 180°, $p < .001$; 90° < 180°, $p < .001$; 90° < 180°, $p = .006$; all other *ps* > .05). In comparison, the main effect of distance in the TD 10- to 11-year-olds and ADHD groups were best explained by the linear contrast (as outlined above). Overall, the two youngest age groups show a flatter linear increase in RT and children with ADHD show a pattern akin to the TD 10 to 11-year-olds and thus show an age-appropriate pattern. Mean and standard deviations of raw mental rotation RTs are reported in the **Appendix** (Table A.2.1.1).

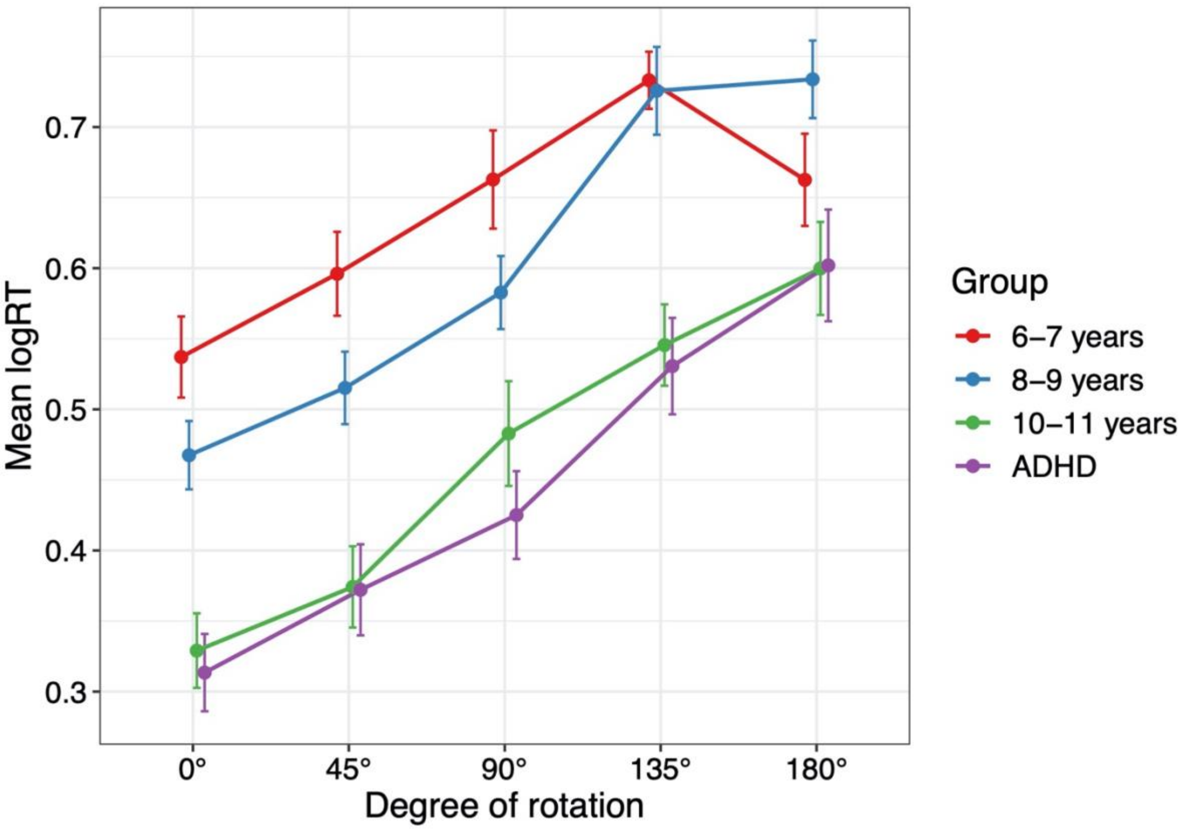



*Figure 3.2:* Means and SE of logRT for each degree of rotation (0°, 45°, 90°, 135°, 180°) in the mental rotation task

Lines of best fit determined slope value, i.e., the gradient of the line, of mean correct RT across each degree of rotation (0°, 45°, 90°, 135°, 180°) for each participant. A one-way ANOVA of slope values was conducted to examine how slopes differed by group. There was a significant main effect of group (*Welch's* $F(3.00,44.38) = 5.94$, $p = .002$, $\eta^2 = .16$). Post hoc comparisons revealed a significantly steeper slope between TD 8- to 9-year-olds ($M = .74$, $SD = .45$) than the ADHD group ($M = .32$, $SD = .19$). All other differences were not significant (TD 6- to 7-year-olds: $M = .39$, $SD = .42$; TD 10- to 11-year-olds: $M = .52$, $SD = .29$; $ps > .05$). Thus, children with ADHD show the same pattern as the TD group of the same chronological age.

3.3.4.2. Mental rotation accuracy

An equivalent ANOVA was conducted on mental rotation percentage accuracy per degree of rotation by group. There was a significant main effect of group ($F(3,84) = 4.66$, $p = .005$, $\eta_p^2 = .14$), where post hoc comparisons showed significantly greater accuracy in the ADHD group compared to TD 6-7-year-olds ($p$ = .011) and TD 8-9-year-olds ($p$ = .016) (all other $ps$ > .05). There was a significant main effect of degree of rotation ($F$(1.97,165.40) = 29.76, $p$ < .001, $\eta_p^2$ = .26), which is best explained by the significant interaction between degree of rotation and group ($F$(5.91,165.40) = 2.56, $p$ = .022, $\eta_p^2$ = .08) (see Figure 3.3). The main effects of degree of rotation for the TD 6- to 7-year-olds and TD 8- to 9-year-olds were best explained by linear contrasts (TD 6- to 7-year-olds: $F$(1,26) = 26.89, $p$ < .001, $\eta_p^2$ = .51; TD 8- to 9-year-olds: $F$(1,18) = 18.27, $p$ < .001, $\eta_p^2$ = .50), indicative of accuracy decreasing with increasing degree of rotation. The TD 10-11-year-olds demonstrated a significant main effect of degree of rotation ($F$(1.94,42.63) = 4.93, $p$ = .019, $\eta_p^2$ = .17), however pairwise comparisons revealed no significant differences in accuracy between any degree of rotation ($ps$ > .05), indicative of a shallow slope. While this is surprising given the main effect, accuracy was lower at 45° compared to 0°, however Bonferroni corrected alpha was just above .05 ($p$ = .052). The ADHD group showed no significant main effect of degree of rotation ($F$(1.84,33.19) = 2.07, $p$ = .146, $\eta_p^2$ = .10). Taken together, accuracy decreased as degree of rotation increased only in the two youngest TD groups, whilst the TD 10 to 11-year-olds and children with ADHD showed a broadly similar profile. While this implies possible ceiling effects in the TD 10-11-year-olds and ADHD groups, one sample t tests of mean percentage in accuracy for each degree of rotation against a maximum score of 100%, showed accuracy was significantly below from 100% for each degree of rotation for both the TD 10- to 11-year-olds (0°: $t$(22) = -5.88, $p$ < .001, $d$ = -1.22; 45°: $t$(22) = -3.02, $p$ = .006, $d$ = -.63; 90°: $t$(22) = -3.04, $p$ = .006, $d$ = -.63; 135°: $t$(22) = -3.83, $p$ = .001, $d$ = -.79); 180°: $t$(22) = 3.76, $p$ = .001, $d$ = -.78) and the ADHD group (0°: $t$(18) = -3.17, $p$ = .005, $d$ = -.73; 45°: $t$(18) = -3.89, $p$ = .001, $d$ = -.89; 90°: $t$(18) = -2.48, $p$ = .023, $d$ = =.57; 135°: $t$(18) = -2.17, $p$ =.043, $d$ = -.49; 180°: $t$(18) = -2.79, $p$ = .012, $d$ = .64). This demonstrates accuracy was not at ceiling for either group.

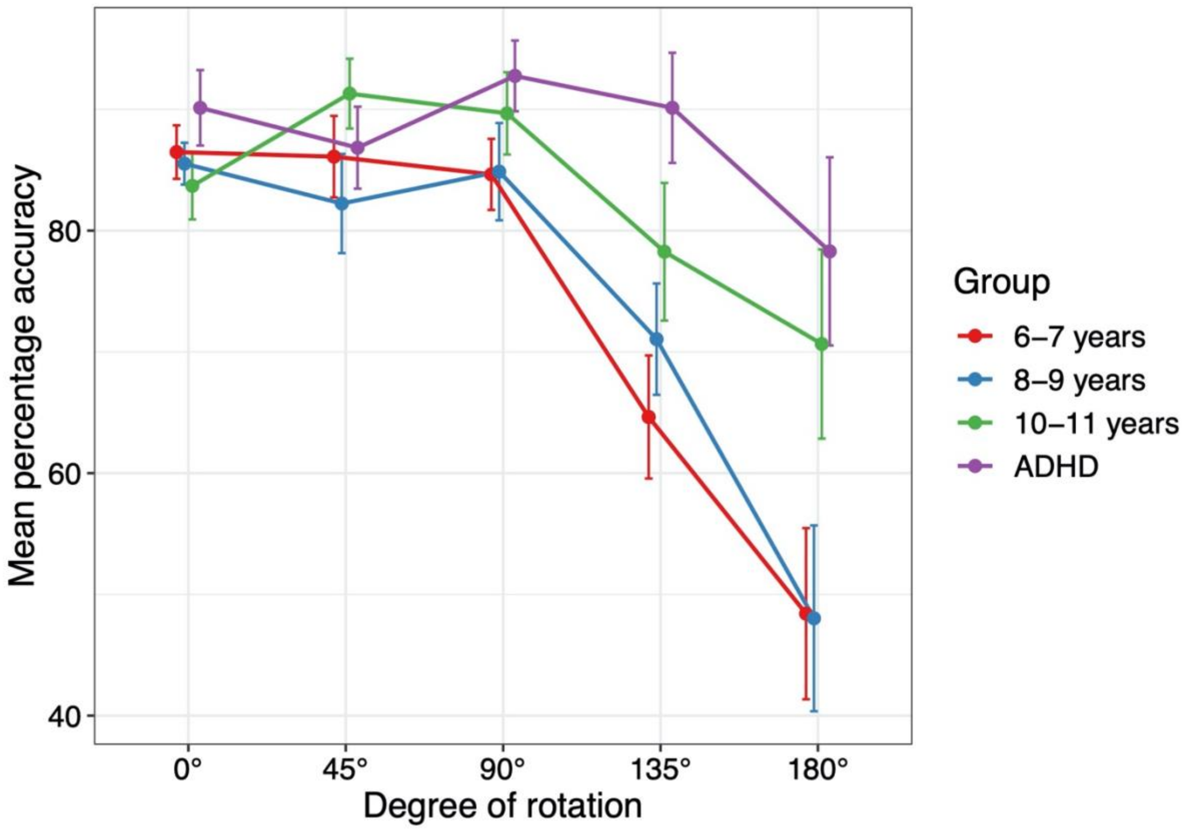



*Figure 3.3:* Means and SE of percentage accuracy for each degree of rotation (0°, 45°, 90°, 135°, 180°) per group in the mental rotation task

**3.3.5. Image scanning**

A mixed ANOVA was conducted of mean logRT with within-subject factor of distance (70mm, 81mm, 100mm, 154mm, 262mm) and between-subject factor of group. There was no main effect of group ($F(3,84) = 1.07$, $p = .367$, $\eta_p^2 = .04$), alongside a significant main effect of distance ($F(3.34,280.91) = 56.76$, $p < .001$, $\eta_p^2 = .40$). Post hoc comparisons revealed that there were no significant differences in logRT between the shortest distance and the next two shortest distances, 70mm and 81mm ($p = .214$), 70mm and 100mm ($p = 1.00$), 81mm and 100mm ($p = .128$), alongside significant increases in logRT between all other distance pairs (*ps* < .001). There was also a significant interaction between distance and group ($F(10.03,280.91) = 1.94$, $p = .040$, $\eta_p^2 = .06$). Follow up RM ANOVAs were conducted to interpret the interaction. All groups showed a significant main effect of distance (TD 6-7-year-olds: $F(3.00,78.00) = 14.52$, $p < .001$, $\eta_p^2 = .55$; TD 8-9-year-olds: $F(2,79,50.30) = 6.88$, $p < .001$, $\eta_p^2 = .28$; TD 10-11-year-olds: $F(3.39,74.69) = 32.24$, $p < .001$, $\eta_p^2 = .59$; ADHD group: $F(2.94,52.98) = 16.18$, $p < .001$, $\eta_p^2 = .47$), however slightly different patterns of logRT were found in each group. The TD 6- to 7-year-olds demonstrated significant increase in logRT between each distance and the furthest distance only (*ps* < .001). The TD 8- to 9-year-olds demonstrated significantly longer RTs between the two shortest distances and the two furthest distances only (*ps* < .05). The TD 10- to 11-year-olds exhibited a similar pattern of results to the mixed ANOVA in that logRTs did not increase between the shortest distance and next two shortest distances (*p*s < .05) but did increase between all other distance pairs (*p*s < .001) apart from between the two furthest distances (154mm and 262mm, *p* = .224). Finally, the ADHD group showed the same pattern of results are the 10- to 11-year-olds (*p*s < .05). Thus, the younger children exhibit flatter slopes than older children, and children with ADHD (Figure 3.4). Mean and standard deviations of raw image scanning RTs are reported in the **Appendix** (Table A.2.1.2).

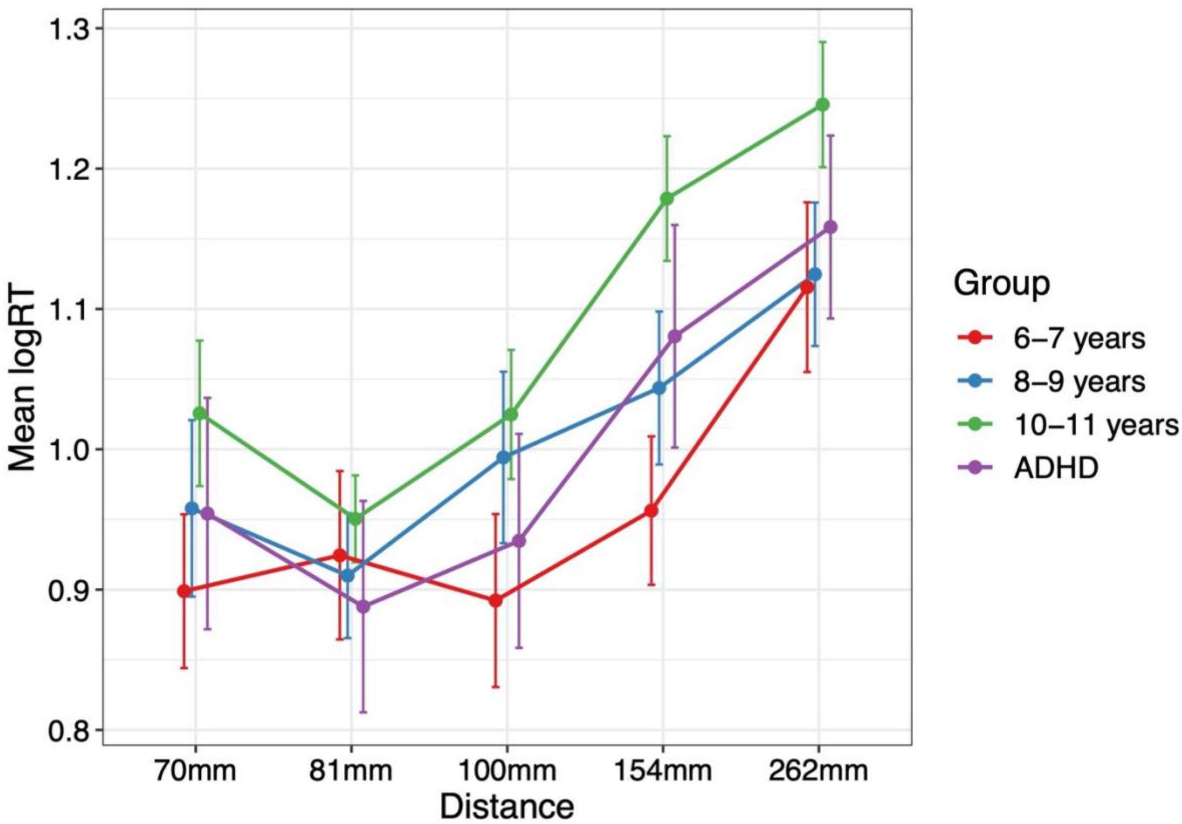



*Figure 3.4:* Means and SE of logRT for each distance (70mm, 81mm, 100mm, 154mm, 262mm) per group in the image scanning task

Slope values were calculated for each participant on the raw RT data and a one-way ANOVA examining group differences in slope values was conducted. There was a significant main effect of group (*Welch's* $F(3.00,45.61) = 8.55$, $p < .001$, $\eta^2 = .22$). Post hoc comparisons showed significantly shallower slopes between the ADHD group ($M = .01$, $SD = 1.46$) and the TD 6- to 7-year-olds ($M = 1.39$, $SD = 1.84$) ($p = .011$) and between the ADHD group and the TD 10- to 11-year-olds ($M = 2.16$, $SD = 1.22$) ($p < .001$) but the difference in steepness of slope was not significant between the ADHD group and TD 8- to 9-year-olds ($M = 1.19$, $SD = 1.03$) ($p = .081$). Thus, children with ADHD appear to show a different pattern of results to TD children of the same age.

An equivalent mixed ANOVA was conducted on perception control trial logRT with distance (70mm, 81mm, 100mm, 154mm, 262mm) as the within-subject factor and group as the between-subject factor. This was conducted to establish whether participants demonstrated similar patterns of results in shifting attention between varying distances on the screen (perception control trials) compared to shifting attention between varying distances in mind in the absence of sensory input (image scanning trials), and whether this pattern of results is the same in TD children compared to ADHD. As shown in **Chapter 2**, a similar pattern to the MI scanning condition was found in the perception control trials. First, there was no main effect of group ($F(3,84) = 1.04$, $p = .379$, $\eta_p^2 = .04$), but there was a significant main effect of distance ($F(3.18,336.00) = 78.02$, $p < .001$), $\eta_p^2 = .48$). This suggested as distance increases, logRT increased and pairwise comparisons revealed the same relationships between each distance as in the MI condition reported above (*ps* < .001), except for a significant increase in logRT between 70mm and 81mm ($p = .038$). There was also a significant interaction between distance and group ($F(12.00,336.00) = 2.457$, $p = .004$, $\eta_p^2 = .08$). Follow up RM ANOVAs were conducted to interpret the interaction. Firstly, as above, all groups showed a significant main effect of distance (TD 6- to 7-year-olds: $F(2.23,58.00) = 23.20$, $p < .001$, $\eta_p^2 = .47$; TD 8- to 9-year-olds: $F(3.18,57.29) = 11.94$, $p < .001$, $\eta_p^2 = .39$; TD 10- to 11-year-olds: $F(2.30,50.69) = 27.30$, $p < .001$, $\eta_p^2 = .55$; ADHD group: $F(2.57,46.16) = 31.99$, $p < .001$, $\eta_p^2 = .63$). The TD 6- to 7-year-olds demonstrated differences in logRT similar to the mixed ANOVA in that there were no significant differences in logRT between the three shortest distances alongside significant increases in logRT between all other shorter and further distance pairs (*ps* < .001). The other TD groups and the ADHD group demonstrated the same differences between distances as the TD 6- to 7-year-old group except for no significant increase in logRT between 154mm and 262mm (*ps* > .05). Thus, all TD children and the ADHD sample demonstrated a broadly similar pattern of results in perception control trials compared to image scanning trials (Figure 3.5). Mean and standard deviations of raw perception control trial RTs are reported in the **Appendix** (Table A.2.1.3).

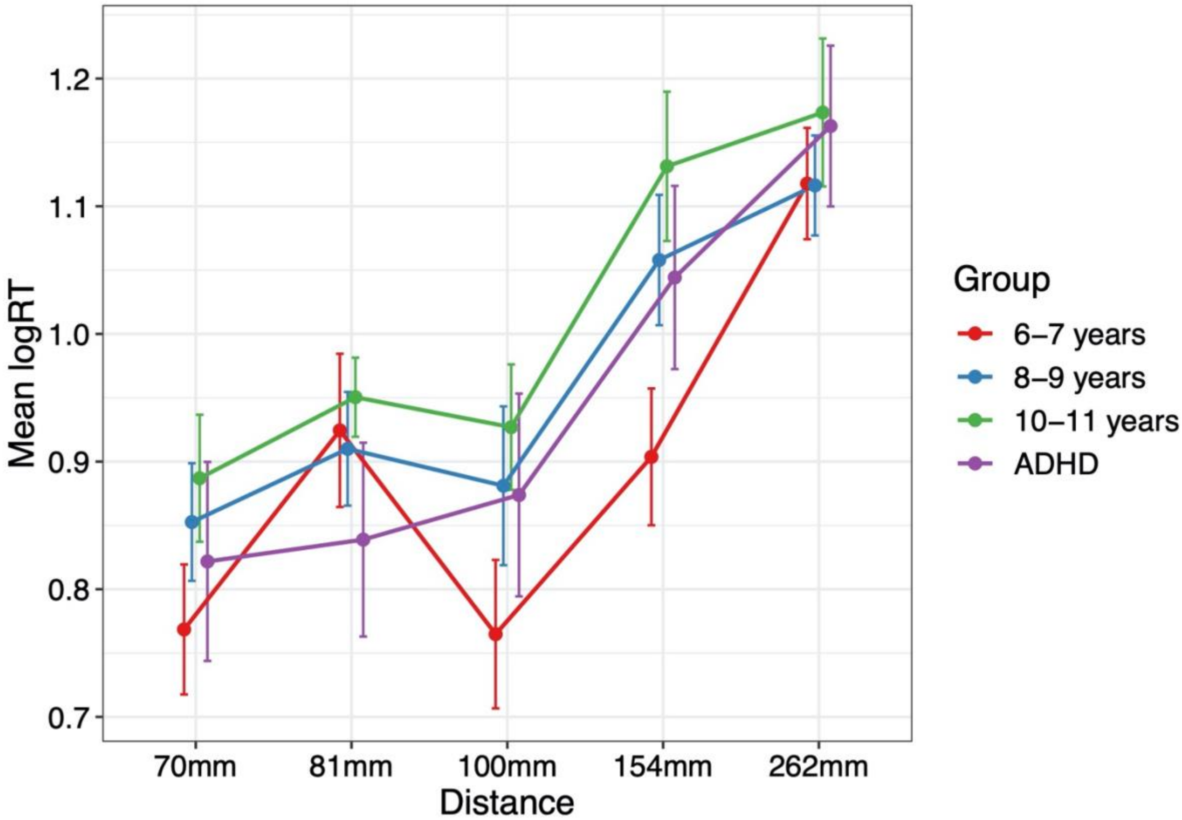



*Figure 3.5:* Means and SE of logRT for each distance (70mm, 81mm, 100mm, 154mm, 262mm) per group in perception control trials of the image scanning task

3.3.5.1 Evaluating the ability of children with ADHD to internally represent distance in image scanning

Actual distance ratios (ratio 1: 70mm to 81mm = 1.16, ratio 2: 70mm to 100mm = 1.43, ratio 3: 70mm to 154mm = 2.2, ratio 4: 70mm to 262mm = 3.74) and imaged distance ratios (derived from RT, e.g., ratio 1 = RT ratio between 70mm to 81mm) were calculated as in **Chapter 2.** One sample t tests between each actual ratio and each imaged distance ratio were then conducted. Notably, this revealed that children with ADHD significantly underestimate distance, as do their typical peers (ratio 1: $t(18) = -3.19$, $p = .005$, $d = -.73$; ratio 2: $t(18) = -5.48$, $p < .001$, $d = -1.26$; ratio 3: $t(18) = -8.19$, $p < .001$, $d = -1.88$; ratio 4: $t(18) = -9.91$, $p < .001$, $d = -2.27$). As can be seen in Figure 3.6, underestimation is greater at further distances. A mixed ANOVA was conducted with a within-subject factor of difference scores (actual distance ratios subtracted from imaged distance ratios) and between-subject factor of group to test this observation statistically and to examine whether this effect was present in all groups. There was no significant main effect of group ($F < 1$). There was a significant main effect of difference score which was best explained by a significant linear contrast ($F(1.83,153.89) = 343.75$, $p < .001$, $\eta_p^2 = .80$). Post hoc comparisons revealed significant increases between each difference score (*ps* < .001). This shows that as distance increased, underestimation of distance significantly increased. There was no significant interaction between difference scores and group ($F(5.49,153.89) = 1.51$, $p = .183$, $\eta_p^2 = .05$). Overall, this suggests children with ADHD and TD children across the age bands show the same pattern of results in the image scanning task. The equivalent analyses were conducted on perception control trials and the same relationships were indicated (see **Appendix A.2.2.**).

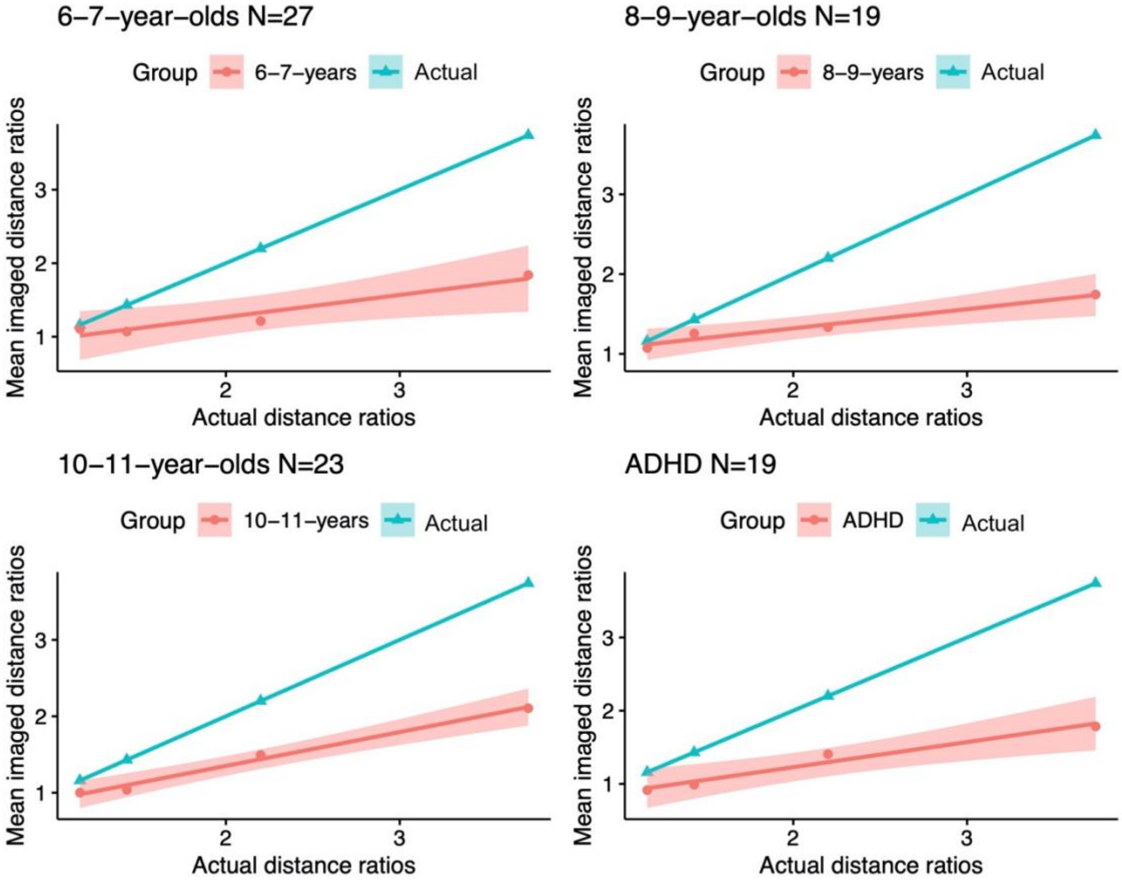



*Figure 3.6:* Scatterplots depicting the relationship between actual distance ratios and mean imaged distance ratios with confidence intervals and best fitting regression lines for each group. Both actual distance ratios and imaged distance ratios are plotted for reference

### 3.3.6. VWM ability

To establish how children with ADHD performed in VWM maintenance and VWM manipulation compared to TD children, a mixed ANOVA on sequence length was conducted with the within-subject factor of VWM task (forward span/maintenance, backward span/manipulation) and between-subject factor of group. Sequence length is comprised of the longest sequence length participants correctly remembered in the forward and backward span tasks, respectively. First, there was a significant main effect of group ($F(3,84) = 4.61$, $p = .005$, $\eta_p^2 = .14$) where post hoc comparisons revealed significantly lower performance in the TD 6- to 7-year-olds compared to the TD 10- to 11-year-olds ($p = .007$) (all other *ps* > .05). The main effect of task was not significant

($F(1,84) = 2.23$, $p = .139$, $\eta_p^2 = .03$) and there was no significant interaction between VWM task and group ($F < 1$) (see Figure 3.7). Thus, children with ADHD demonstrated VWM abilities in line with TD children of the same chronological age and the same pattern of abilities as TD controls.

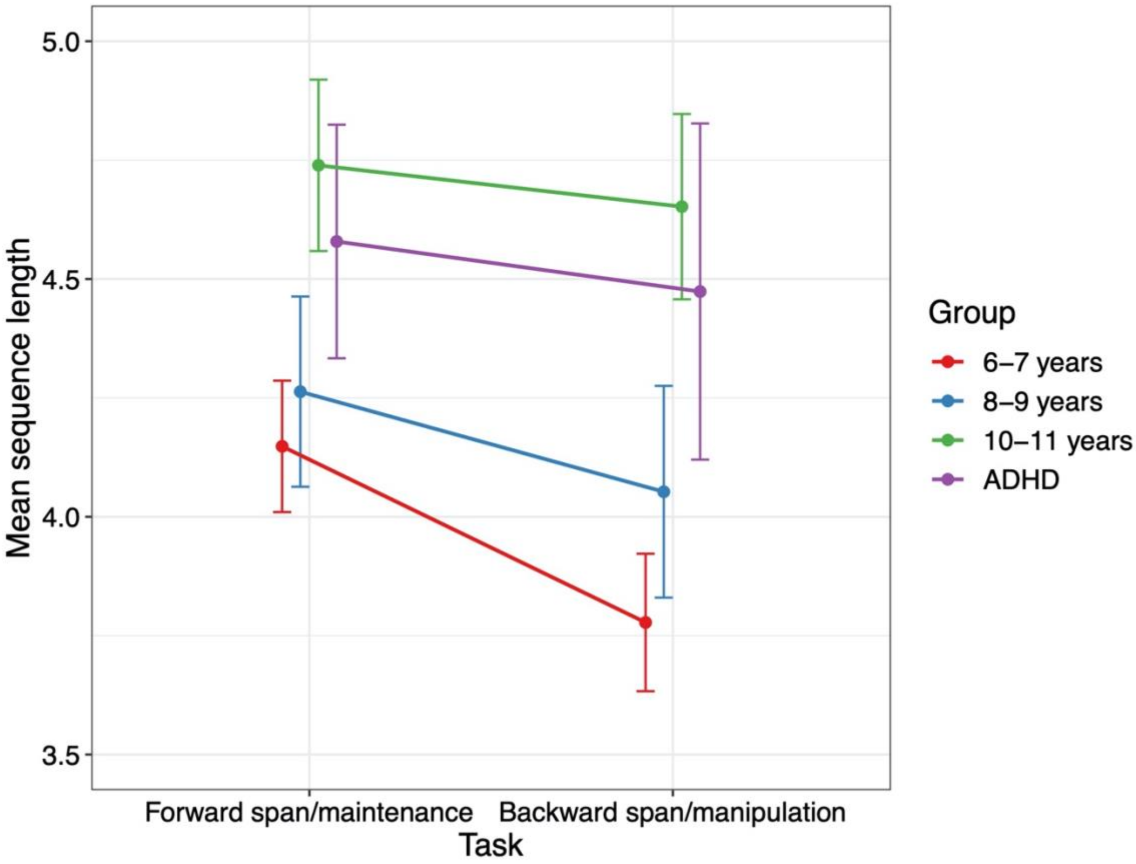



*Figure 3.7:* Means and SE of forward span/maintenance and backward span/manipulation sequence length per group in the VWM tasks

### 3.3.7. Examining the profile of MI and VWM abilities in ADHD

To examine profiles of ability in each of the components of MI and VWM in the ADHD group, Z-scores were computed based on the mean and SD of the TD group that matched at a group level on chronological age (10- to 11-year-olds) to the ADHD group. Z-scores were computed for each of the accuracy measures: image generation overall accuracy, image maintenance overall accuracy, mental rotation accuracy and image scanning index of metric properties (IMP), VWM maintenance sequence length, VWM manipulation sequence length. The image generation and image maintenance overall accuracy scores were derived as in **Chapter 2**, whereby high precision trials were scored as 2, low precision trials and correct same trials were scored as 1 and incorrect trials scored as 0. This was summed to form an overall accuracy score (max score = 25). The image scanning IMP score is the mean squared error of imaged distance RT ratios, which was also calculated as in **Chapter 2**, where a score of 0 indicates imaged distance ratios mapped exactly to actual distance ratios (6- to 7-year-olds: *M* = 1.51, *SD* = .69, *range* = .30 to 3.24; 8- to 9-year-olds: *M* = 1.56, *SD* = .81, *range* = .38 to 2.72; 10- to 11-year-olds: *M* = 1.34, *SD* = .71, *range* = .06 to 2.84; ADHD group: *M* = .89, *SD* = .44, *range* = .09 to 1.67). A mixed ANOVA was conducted on Z-scores with task (image generation, image maintenance, mental rotation, image scanning, VWM maintenance, VWM manipulation) as the within-subject factor and group (TD 10-11-year-olds, ADHD) as the between-subject factor. There was no main effect of group ($F(1,40) = 1.30$, $p = .260$, $\eta_p^2 = .03$) and no significant interaction between task and group ($F < 1$) (see Figure 3.8). Thus, this group of children with ADHD are performing at an age-appropriate level in all measures and show a typical profile of performance.

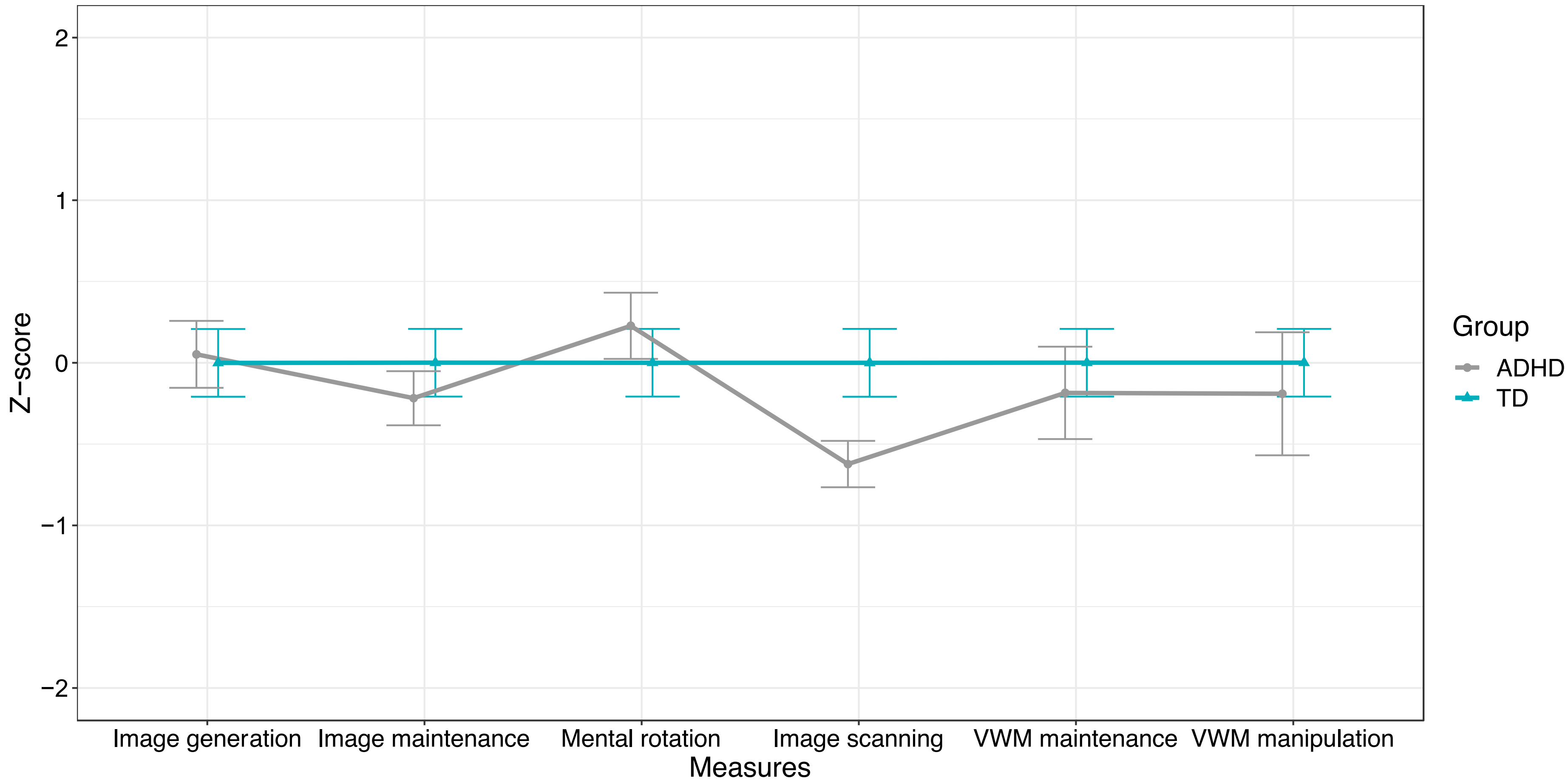



*Figure 3.8:* Mean and SE of Z-scores for each component of MI (image generation overall accuracy, image maintenance overall accuracy, mental rotation accuracy and image scanning IMP) and VWM (maintenance and manipulation) per group

### 3.3.8. Using Latent Profile Analysis to investigate individual differences in MI abilities beyond diagnostic labels

To investigate how the MI and VWM components map onto distinct profiles of scores, a latent profile analysis was conducted across the whole sample including both TD children and children with ADHD (N = 88). The variables included in the group differences analysis were also included here; image generation accuracy, image maintenance accuracy, mental rotation accuracy, image scanning IMP, VWM maintenance and VWM manipulation. The criteria applied to select the best model was as follows: 1) low Bayesian Information Criteria (BIC) and sample-adjusted BIC (SABIC) indicating less unexplained variance and better model fit, 2) significant Bootstrap Likelihood Ratio Test (BLRT) indicating a statistically significant differences between *k* profiles and *k-1* profiles, 3) entropy value above .90 where a value closer to 1 indicates higher probability of accurate class separation.

3.3.7.1. Model fit

To determine the optimum number of profiles, the model fit of 1:6 profiles was assessed (see Table 3.2 for model fit statistics). While entropy values are lower than .90, the highest values apart from the 1-profile model are indicated in the 5-profile and 6-profile models. The BLRT *p* values indicate a significant difference between the 3- and 2-profile models as well as between the 5- and 4-profile models. There is a significant difference between the 4-profile and 5-profile models (*p* = .009) and the 5-profile model has a greater entropy value (.79) and the lowest SABIC value (2295.69), which suggests the 5-profile model is the best fit for the data. Entropy values, SABIC and BIC values were not dissimilar between the 6-profile model and 5-profile model, however there was no significant difference between the 5-profile and 6-profile models, therefore the 5-profile model was chosen as the final model.

Table 3.2

*Model fit statistics for latent profile analysis*

| **Model** | **BIC** | **SABIC** | **Entropy** | **BLRT p value** | **N assigned to each Profile (P)** |
|---|---|---|---|---|---|
| 1-profile | 2367.02 | 2329.15 | 1 | - | P1=88 |
| 2-profile | 2384.08 | 2324.12 | 0.49 | $p$ = .15 | P1=41, P2=47 |
| 3-profile | 2387.20 | 2305.16 | 0.73 | $p$ = .019 | P1=13, P2=37, P3=38 |
| 4-profile | 2420.07 | 2315.93 | 0.76 | $p$ = .960 | P1=36, P2=28, P3=5, P4=19 |
| 5-profile | 2421.91 | 2295.69 | 0.79 | $p$ = .009 | P1=14, P2=16, P3=11, P4=20, P5=27 |
| 6- profile | 2443.99 | 2295.68 | 0.81 | $p$ = .603 | P1=5, P2=16, P3=9, P4=9, P5=23, P6=26 |

### 3.3.7.2. Transdiagnostic profiles of MI and VWM abilities

Means and standard deviations of raw scores per variable for each profile are presented in Table 3.3 and means and SE of Z-scores for each profile are presented in Figure 3.9. Profiles are described as a comparison of abilities relative to the other profiles. Profile 1 included participants with relatively high mental rotation abilities, VWM maintenance and VWM manipulation abilities and moderate abilities in all other measures. Profile 2 included participants with relatively low mental rotation abilities alongside moderate abilities in other measures. Profile 3 included participants with relatively low image generation, image maintenance and VWM maintenance abilities and moderate abilities in all other measures. Profile 4 include participants with relatively high mental rotation abilities and moderate abilities in all other measures. Finally, profile 5 included participants with moderate abilities in all MI and VWM measures, relative to the other profiles.

Table 3.3.

*Means and standard deviations of score per variable for each profile*

| Measures | Profile 1 N=14 | Profile 2 N=16 | Profile 3 N=11 | Profile 4 N=20 | Profile 5 N=27 |
|---|---|---|---|---|---|
| | | | Mean (SD) | | |
| Image generation | 12.79 (3.40) | 14.69 (3.55) | 10.91 (2.12) | 15.30 (2.08) | 15.93 (3.08) |
| Image maintenance | 15.21 (2.86) | 14.38 (2.68) | 11.55 (2.46) | 14.80 (2.21) | 17.78 (2.01) |
| Mental rotation | 91.25 (5.94) | 57.5 (6.58) | 78.18 (10.84) | 95.63 (4.13) | 77.99 (8.11) |
| Image scanning | 1.27 (.53) | 1.46 (.75) | 1.56 (.95) | .85 (.49) | 1.60 (.66) |
| VWM maintenance | 5.64 (.49) | 4.81 (.54) | 3.36 (.50) | 3.90 (.72) | 4.37 (.56) |
| VWM manipulation | 5.29 (.99) | 4.19 (.91) | 3.73 (.90) | 4.30 (1.03) | 3.81 (1.04) |

One-way ANOVAs were conducted with each MI and VWM measure as the dependent variables and profile as the fixed factor to test whether the profiles exhibited statistically different levels of ability. Each one-way ANOVA revealed a significant main effect of profile (image generation: *Welch's* $F(4.00,35.67) = 10.25$, $p < .001$, $\eta^2 = .26$; image maintenance: *Welch's* $F(4.00,34.12) = 16.01$, $p < .001$, $\eta^2 = .42$; mental rotation: *Welch's* $F(4.00,33.99) = 106.51$, $p < .001$, $\eta^2 = .78$; image scanning: *Welch's* $F(4.00,34.46) = 5.59$, $p = .001$, $\eta^2 = .17$; VWM maintenance: $F(4.00,36.25) = 35.64$, $p < .001$, $\eta^2 = .59$; VWM manipulation: $F(4.00,36.05) = 5.59$, $p = .001$, $\eta^2 = .22$). Post hoc comparisons revealed the relative differences between profiles described above were supported. Image generation accuracy was significantly lower in profile 3 compared to profile 2 ($p = .016$), profile 4 ($p < .001$) and profile 5 ($p < .001$) (all other comparisons: *p*s > .05). Image maintenance accuracy was also significantly lower in profile 3 compared to profile 1 ($p = .017$), profile 4 ($p = .013$) and profile 5 ($p < .001$) and there was a trend for lower accuracy between profile 3 and profile 2 ($p = .066$). Image maintenance accuracy was also significantly lower in profile 4 compared to profile 5 ($p < .001$). Mental rotation accuracy in profile 1 was significantly greater than all other profiles ($p < .02$) except for profile 4 ($p = .159$). All profiles indicated significantly greater accuracy than profile 2 (*p*s < .001). Accuracy in profile 3 was also significantly lower than all other profiles except for profile 5 ($p = 1.00$). Finally, profile 4 indicated greater accuracy than profile 5 ($p < .001$). There were no significant differences in image scanning apart from significantly lower IMP score (and thus significantly lower deviation in imaged distance compared to actual distance) in profile 4 compared to profile 5 ($p < .001$). In VWM maintenance, profile 1 indicated the greatest performance in that sequence length was significantly higher in this profile than all other profiles (*p*s < .001). Profile 2 demonstrated significantly greater performance than profile 3 and profile 4 (*p*s < .001). Profile 3 indicated significantly poorer performance than profile 5 ($p < .001$) (all other comparisons: *p*s > .05). Finally, profile 1 showed greater performance in VWM manipulation compared to all other profiles (*p*s < .05; trend for profile 4: $p = .06$). All other comparisons between profile performance in VWM manipulation were not significant (*ps* > .10).

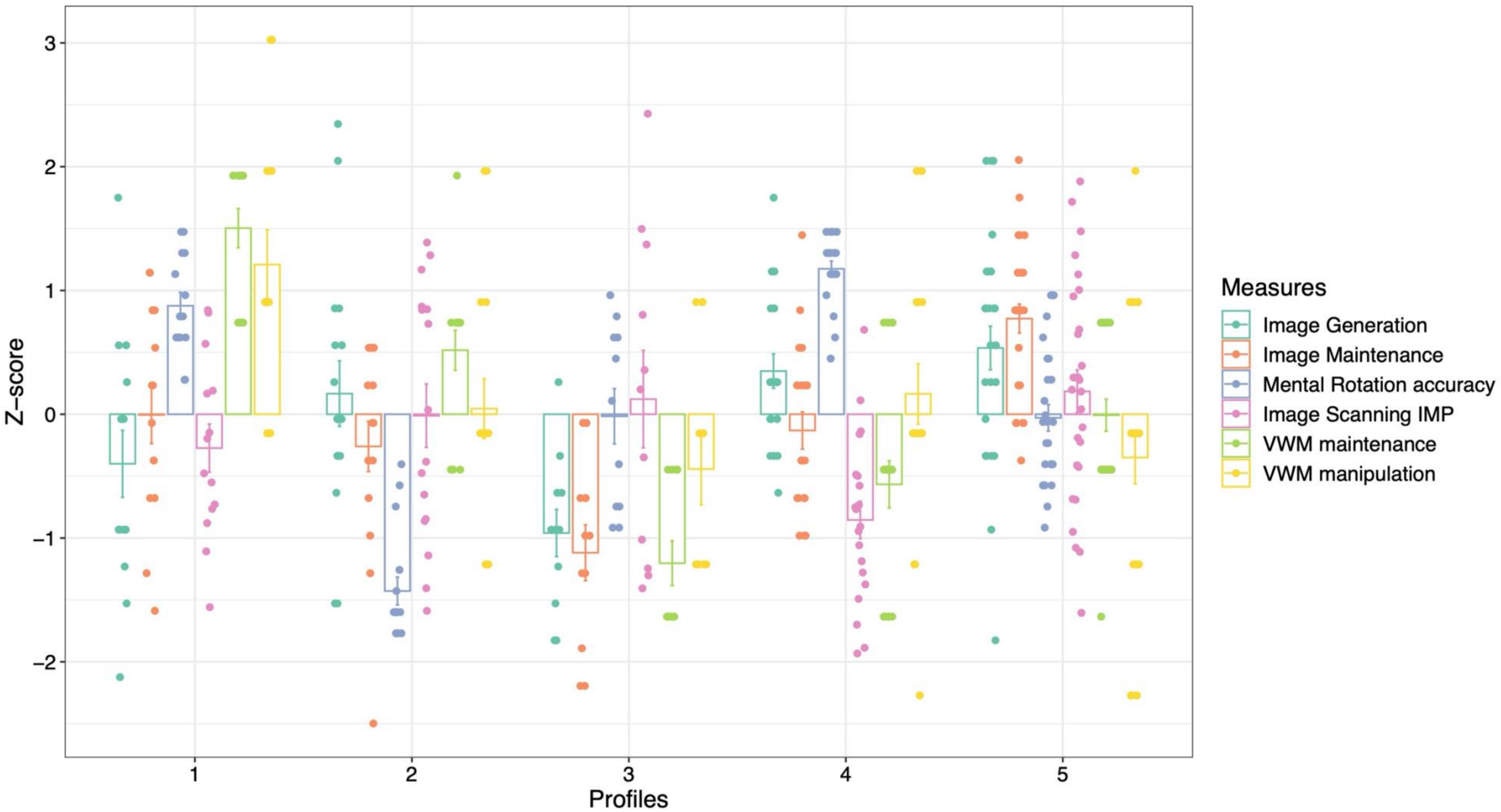



*Figure 3.9:* Mean and SE of Z-score and individual data points for each measure per profile

Next, profiles were examined to see if they reflect diagnostic groups (see Figure 3.10). The profiles were transdiagnostic in that they included both TD children and children with ADHD, with the exception of profile 3 which included only children from the TD group. Accordingly, children with ADHD made up 36% (N = 5) of profile 1 (relatively high mental rotation, VWM maintenance and VWM manipulation abilities) and children with ADHD made up 13% (N = 2) of profile 2 (relatively low mental rotation ability). Only TD children were included in profile 3 (relatively low image generation, image maintenance and VWM maintenance abilities). There was a fairly even split of TD children in profile 4 (relatively high mental rotation ability) with children with ADHD making up 40% (N = 8) of this profile. Finally, profile 5 had the largest number of participants and was characterised as relatively moderate abilities in all measures, whereby children with ADHD made up 15% (N = 4) of the participants in this profile.

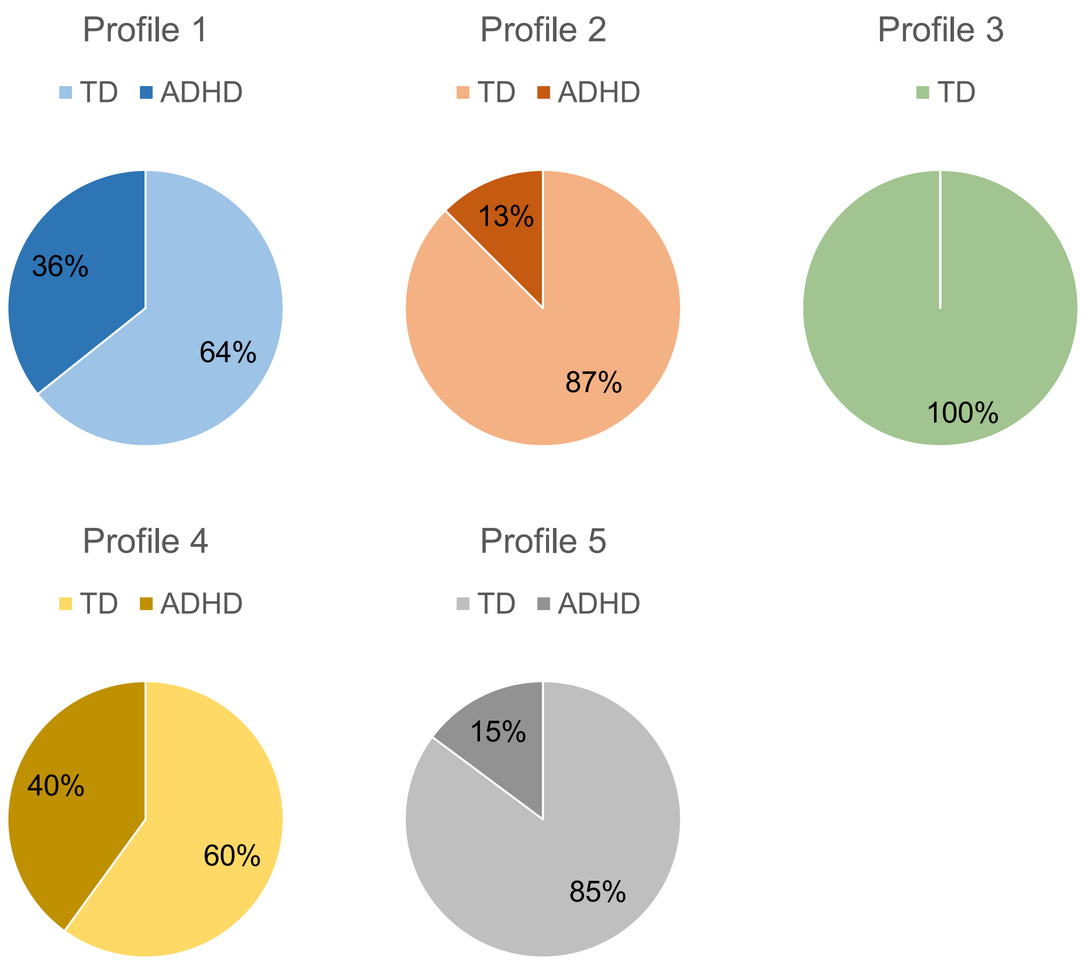



*Figure 3.10:* Pie charts depicting percentage of TD children and children with ADHD in each profile

### 3.3.9. The relationship between each component of MI, VWM and ADHD symptoms

The following variables were included in the ADHD group correlation analyses: image generation overall accuracy, image maintenance overall accuracy, mental rotation correct RT, image scanning IMP, VWM maintenance (forward span), VWM manipulation (backward span) and ADHD symptoms (Conners Global ADHD index T score). As outlined in Table 3.4, there were no significant associations between each of the MI components, supporting a separable component model of MI in children with ADHD (*ps* > .001). There were also no relationships between components of MI and either measure of VWM (*ps* > .001) or between components of MI and ADHD symptoms (*ps* > .001). Overall, ability in the components of MI appear to be dissociated from VWM maintenance and manipulation ability, as well as ADHD symptoms, in children with ADHD.

Table 3.4

*Kendall's Tau correlations (age partialled out) between components of MI, VWM and ADHD symptoms in children with ADHD (N=19)*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. |
|---|---|---|---|---|---|---|---|
| 1. Image generation | — | | | | | | |
| 2. Image maintenance | -.035 | — | | | | | |
| 3. Mental rotation RT | .093 | -.018 | — | | | | |
| 4. Image scanning IMP | .056 | .058 | .177 | — | | | |
| 5. Forward VWM | -.079 | .301 | -.108 | .085 | — | | |
| 6. Backward VWM | .066 | -.020 | -.258 | -.095 | .183 | — | |
| 7. Conners Global ADHD Index | -.016 | -.148 | .070 | .123 | .253 | -.230 | — |

Spearman's correlation analyses were conducted on the TD group controlling for age (Table 3.5). The findings were in line with the findings in the broader sample of children in **Chapter 2** in that there was support for a separable-component model of MI and support for a dissociation between MI components and VWM measures due to a lack of significant associations between the MI components and between MI components and VWM measures (*ps* > .002, multiple comparison cut-off). Moreover, in line with the ADHD group correlation, there were no significant associations between ADHD symptoms and either the MI components or VWM measures (*ps* > .001).

Table 3.5.

*Spearman's correlations on the residuals on the ranks of the variables (age partialled out) between components of MI, VWM and ADHD symptoms in TD children age 6-11 years (N=69)*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. |
|---|---|---|---|---|---|---|---|
| 1. Image generation | — | | | | | | |
| 2. Image maintenance | .264 | — | | | | | |
| 3. Mental rotation RT | -.034 | -.217 | — | | | | |
| 4. Image scanning IMP | -.002 | .135 | -.201 | — | | | |
| 5. Forward VWM | -.001 | .059 | .001 | .098 | — | | |
| 6. Backward VWM | .029 | -.136 | -.045 | -.162 | .164 | — | |
| 7. Conners 3 T score | -.080 | -.059 | -.047 | .138 | -.059 | -.104 | — |

*Note.* Conners 3 T score is comprised of the highest T score on either the Hyperactivity scale or the Inattentive scale on the CTRS-3:L (see **Section 3.2.3.1.** for details)

## 3.4. Discussion

The overarching goal of this chapter was to characterise MI abilities alongside VWM in children with ADHD. To do this, the first aim was to examine abilities in each of the components of MI in children with ADHD compared to TD children in group-level analyses and individual differences analyses. The second aim was to investigate how components of MI are associated with maintenance and manipulation measures of VWM and symptoms of ADHD. With reference to the first aim, between-group analyses revealed children with ADHD perform at a chronological age-appropriate level in all MI tasks and broadly present the same pattern of effects as their age-matched TD peers, except for shallower slopes in the image scanning task. In line with this, it was also found that the ADHD group profile of abilities in MI and VWM was at the same level as the TD children of the same age. Comparatively, data-driven, individual-level analyses (latent profile analysis) found evidence for five distinct profiles of ability and these profiles were transdiagnostic; each profile of varying abilities in MI and VWM comprised both TD and children with ADHD, with the exception of one profile including only TD children. This suggests that individual differences in MI and VWM in children with ADHD are no different to those observed in TD children, therefore an ADHD-specific atypical profile of MI or VWM was not apparent in this study. With reference to the second aim, support for a separable-component model of MI and distinctions between MI and VWM was found in the ADHD group and the TD group, in line with **Chapter 2**. Moreover, there were no associations between ADHD symptoms and components of MI in the TD children or in children with ADHD.

### 3.4.1. Characterising MI abilities in ADHD: evidence for typical patterns of performance and age-appropriate levels of ability

The ADHD group presented a typical pattern of performance in both image generation and image maintenance, in line with predefined hypotheses. However, contrary to hypotheses, the ADHD performed at an age-appropriate level. Namely, there was evidence for faster RTs in high precision trials compared to low precision trials in all groups. This suggests that children with ADHD can generate a visual mental image with high precision in the same way that TD children generate visual mental images of high precision. Similar findings were demonstrated in image maintenance; there were faster RTs in high precision trials for all groups. This supports the hypothesis that in high precision trials participants are referring to a readily available visual image, as opposed to taking more time in searching for a less efficient strategy in trials of low precision and extends evidence presented in **Chapter 2** by demonstrating that children with ADHD present with typical patterns of performance in both image generation and image maintenance. The comparison between image generation and image maintenance ability also revealed that children with ADHD showed the same pattern of results to TD children; there was no difference between RT of high precision responses between the image generation and image maintenance. Contrary to predictions, children with ADHD performed at a similar level to TD children of the same age in both image generation and image maintenance.

Although it was predicted that children with ADHD would present with poorer performance in mental rotation alongside differential patterns of performance in RT measures, here there is evidence for typical levels of ability and typical patterns of performance in mental rotation. Children with ADHD showed the linear time-degree of rotation effect in mental rotation that is evident in TD children in this chapter, in the broader sample in **Chapter 2** and in previous research (Estes, 1998; Frick et al., 2014; Möhring et al., 2016; Wimmer et al., 2017). This suggests children with ADHD approach mental rotation in a typical manner. Interestingly, and in contrast to previous findings, children with ADHD demonstrated a typical level of accuracy in the mental rotation task. While there are few studies that have investigated mental rotation abilities in ADHD, these have suggested poorer performance in the ADHD group compared to TD controls (Silk et al., 2005; Vance et al., 2007; Williams et al., 2013, Jakobson and Kikas, 2007). Therefore, the finding here of good performance in mental rotation is somewhat surprising. While findings should be interpreted with caution given the size of this ADHD sample, it appears that not all children with ADHD present with poorer mental rotation performance. As noted in the introduction, these studies also include very small sample sizes (N range = 7 to 26 in the ADHD group) (Silk et al., 2005; Vance et al., 2007; Williams et al., 2013). Therefore, it is apparent that previous studies, as well as this study, may not have a sufficient sample size to detect the full range of mental rotation abilities in ADHD. Further research with a larger sample size is required to elucidate individual differences in mental rotation abilities in ADHD.

The image scanning task is the first task in this study that children with ADHD demonstrate evidence of atypical performance. The ADHD group presented a pattern of results in the analysis of image scanning slopes that was akin to their younger peers; the slopes were shallower than the TD 10- to 11-year-old children (i.e., same chronological age as the ADHD group), which is indicative of delay. As children with ADHD tend to present with greater RT variability than TD controls (Kofler et al., 2013), it might be speculated that this is dependent on the confound of RT variability given that the image scanning slope values are derived from RT, however this would have disrupted the other RT findings in this study, and this was not the case. Moreover, RT variability was not explicitly measured in this study, therefore further research is required to support this. Despite the difference in slopes in the image scanning task, the

ADHD group showed the same pattern of results as TD children of the same age in that there was a linear time-distance effect demonstrating as distance increased, RT increased. This is in line with the findings with the broader sample of children presented in **Chapter 2** and with previous research in TD children (Wimmer et al., 2016).

The extent to which children with ADHD accurately represent distances in mind was examined in the image scanning task. In short, children with ADHD presented a typical pattern of results as evidenced by greater underestimation of distance with increasing distance and a typical level of ability in that there was no difference between performance in the ADHD group and TD children of the same age. These underestimation effects were also demonstrated in perception control trials, thus supporting the argument presented in **Chapter 2** that individuals underestimate distance in visual mental images as they do in perceived distance estimation (Giovanni et al., 2009; Thurley & Schild, 2018). Notably, the image scanning IMP score ranged from .09 to 1.29 in the ADHD group, whereby a score of 0 would indicate individuals imaged distance ratios mapped exactly to actual distance ratios. This therefore indicates that while there are individual differences in the accuracy at which distance is represented in mind, some children with ADHD demonstrate highly precise visual mental images in the image scanning task. The novel findings presented here extend previous research to indicate that children with ADHD estimate distance in mental images in a similar vein to distance estimation in visual perception and both the pattern of results and level of ability is on par with TD children of the same age.

Examining the ADHD profile of ability in each component of MI and each measure of VWM revealed typical level of performance in all measures. This is unsurprising, given the findings of the individual ANOVAs for each task. However, it is somewhat surprising given the evidence from previous case-control studies suggesting poorer VWM (Martinussen et al., 2005; Martinussen & Tannock, 2005; Simone et al., 2015; Nikolas & Nigg, 2013) and mental rotation (Silk et al., 2005; Vance et al., 2007’ Williams et al., 2013; Kalobson & Kikas, 2007) in children with ADHD compared to TD children. Previous research has shown that both children with ADHD and TD children perform worse in VWM manipulation (backward VWM span) compared to VWM maintenance (forward VWM span) and children with ADHD perform worse that TD children of the same age (Martinussen et al., 2005; Simone et al., 2016). However, in this study, there were no differences between VWM manipulation and VWM maintenance and no differences between performance in the TD groups compared to the ADHD group. Therefore, this ADHD group appear to present with strong VWM abilities compared to those reported in previous research employing case-control designs. Comparatively, typical level of VWM performance in the ADHD group is not entirely surprising considered in the context of studies indicating individual differences in VWM abilities in ADHD. For example, a recent study found children with ADHD accounted for 21% of a high VWM ability group (Campez et al., 2020). Thus, the ADHD sample in this study appear to represent those with ADHD who have good VWM abilities, and it is possible that this is not representative of the ADHD population. Moreover, it is important to keep in mind that this small sample of children with ADHD might represent those with good MI abilities and therefore it would not be appropriate based on this sample size to make general conclusions about MI abilities in the ADHD population.

Strikingly, the individual-level analyses demonstrate transdiagnostic profiles of ability in the form of five distinct profiles. Prior to interpreting the results of the latent profile analysis, it is important to note that findings regarding the distribution of children per profile should be interpreted cautiously given the imbalance of a much smaller sample size of children with ADHD compared to TD children. Future research involving representative sample sizes are required to support the conclusions made regarding heterogeneity in MI and VWM in these groups. The profiles were characterised as relatively high mental rotation ability and moderate abilities in all other measures (profile 1), relatively low mental rotation ability and moderate abilities in all other measures (profile 2), moderate abilities in MI alongside high abilities in VWM maintenance and VWM manipulation (profile 3), relatively low image generation, image maintenance and VWM maintenance abilities and moderate abilities in all other measures (profile 4) and moderate abilities in all MI and VWM measures (profile 5). The distinctions in profiles of abilities were confirmed by statistical analyses. Importantly, the profiles did not align with diagnostic groups; all profiles included both TD children and children with ADHD, with the exception of profile 3 that included only TD children. This demonstrates the importance of looking beyond diagnostic groups when examining neuropsychological profiles of TD children and children with ADHD.

Considering the profiles in more detail, it is interesting that children with relatively high VWM abilities presented with high mental rotation ability (profile 1), children with relatively low VWM maintenance ability also presented with moderate mental rotation ability (profile 4) and children presenting with relatively low mental rotation ability also presented with moderate VWM abilities (profile 1). These findings appear to be contradictory, however the findings firstly provide further evidence to suggest that components of MI and VWM are separable and secondly, the findings imply individual differences in how children rely on visual information during mental rotation. Adult literature suggests that mental rotation reflects spatial transformation of retained visual information (Hyun & Luck, 2007; Prime & Jolicoeur, 2010) and this is supported by an association between mental rotation and image maintenance in adults in **Chapter 2**. However, in children, it might be the case that children with high VWM abilities (profile 1) or moderate VWM abilities (profile 4) recruit visual information to support mental rotation, whereas those with low VWM maintenance ability (profile 4) do not. This would also be in line with individual differences found in the recruitment of visual strategies in VWM in adults; namely, only “good imagers” appear to use visual strategies in VWM (Keogh & Pearson, 2011; 2014). It is also notable that only 2/16 individuals in the low mental rotation ability profile were children with ADHD, thus seemingly contrasting previous evidence for impaired mental rotation in ADHD (e.g., Williams et al., 2013). However, this supports the above argument that the range of mental rotation abilities in ADHD remains to be fully understood.

It is further notable that only eleven children presented with poor image generation, image maintenance and VWM maintenance abilities and all were TD children. This combined impairment, although in a small subgroup of the sample, implies that while the three abilities are separable in children as a whole (as presented in **Chapter 2** findings and the current study), children that struggle to generate and maintain visual mental images also struggle to maintain visual information in VWM. Such a conclusion would require further analysis of MI abilities in children with poor VWM compared to children with good VWM, which is beyond the scope of the data presented in this study. Finally, the profile indicating moderate abilities in all measures also contained the largest number of participants (N = 28), however only 14% (N = 4) of this group were children with ADHD. This suggests that children with ADHD demonstrated more individual variation in that they were more often grouped into a profile of varying abilities in MI and VWM rather than a profile of moderate abilities in all measures. Taken together, while findings should be interpreted cautiously due to a small sample of children with ADHD, the evidence clearly demonstrates the importance of examining individual differences in both MI and VWM abilities in typical and atypically developing children. Specifically, the data suggests that individual differences observed in MI and VWM are not ADHD-specific and neither atypical profiles of MI nor VWM were identified in the ADHD group. This has implications for the way in which the neuropsychological profile is defined in ADHD and supports the call for adopting transdiagnostic approaches in defining cognitive mechanisms of neurodevelopment (Astle et al., 2021).

### 3.4.2. Evidence for dissociable components of MI and VWM

Evidence for a separable-component model of MI is supported in children with ADHD and in TD children in that there were no significant associations between each of the MI components in either the ADHD group or the TD group. This further supports the notion that children with ADHD appear to present a typical pattern of MI abilities. Secondly, there were no significant correlations between measures of VWM and the components of MI in either group, or between ADHD symptoms and the components of MI or VWM in either the ADHD group or the TD group. First, it is important to note the limitation that the symptom scale measures for the ADHD group and TD group comprised different informants and different scales. Previous literature has suggested low to moderate agreement between parent and teacher ratings (Murray et al., 2018), therefore findings should be interpreted with caution. Moreover, previous literature suggests children with ADHD of different subtypes (e.g., predominantly inattentive or combined) do not differ in their working memory abilities (Alloway et al., 2010). Thus, while there were no predefined hypotheses regarding subtype differences in MI abilities in the ADHD group, further research with suitable sample sizes per ADHD subtype are required to conclude that MI abilities do not differ significantly across ADHD subtypes. Nevertheless, the correlational findings support the argument presented above, and in **Chapter 2,** that while there is previous evidence for shared mechanisms between MI and VWM, the evidence presented in this thesis implies that there is extensive variability in the recruitment of visual strategies in VWM. The absence of associations between MI and VWM in this study is supported by the distinctions between transdiagnostic profiles of ability in that some children present with moderate image generation and image maintenance abilities alongside high VWM maintenance (profile 4), and others present with low image generation and image maintenance abilities alongside low VWM maintenance (profile 3). Therefore, the recruitment of visual strategies in VWM may be dependent on MI ability. Research examining how variability in MI strategies impacts VWM performance is required in both children and adults; **Chapter 4** makes the first step in characterising these individual differences by directly examining this relationship in adults. Finally, the lack of evidence for an association between ADHD symptoms and MI is in line with the evidence presented in this study for typical abilities in MI in children with ADHD.

### 3.4.3. Conclusion

Contrary to predefined hypothesises, children with ADHD demonstrated typical patterns of performance at an age-appropriate TD level in image generation, image maintenance and mental rotation. While the linear time-distance effect in image scanning was found in both TD groups and the ADHD group, children with ADHD demonstrated a shallower slope in image scanning RT compared to TD children of the same age. These findings are notable given that the few studies that have assessed mental rotation abilities in ADHD have suggested poorer performance compared to TD controls. However, evidence for good mental rotation abilities in children with ADHD in this study suggests a broader range of mental rotation ability in children with ADHD than suggested in previous evidence. Moreover, studies with small sample sizes, including this study, likely do not capture the full range of mental rotation abilities in children with ADHD and further research with a representative sample size is required to establish how mental rotation presents in children with ADHD. Another principal finding in this study is the suggestion that individual differences in MI and VWM go beyond diagnostic labels and are not specific to children with ADHD. This is seemingly contrary to previous case-control studies of VWM and instead supports evidence for individual variation in VWM in both children with ADHD and TD children. Importantly, the individual-level analysis also extended the findings in **Chapter 2** by revealing distinct profiles characterising varying abilities in MI and VWM performance across participants. This can be considered alongside evidence for a separable-component model of MI and evidence for distinct abilities in MI compared to VWM maintenance and manipulation in both TD children and children with ADHD. While this supports the notion of dissociated abilities, the individual-level analyses suggest MI and VWM may not be entirely dissociable in children. Instead, the way in which children recruit visual strategies in VWM is likely dependent on varying VWM ability, however further research with children with a broader range of VWM abilities is required to clarify this interpretation. The novel findings presented in this chapter have important implications for understanding how visual mental images are recruited in children's thinking, memory and learning in TD and ADHD. **Chapter 4** will make the first step in deciphering these implications by examining how MI is recruited within a VWM task. Overall, this chapter presents the first known characterisation of MI abilities alongside VWM abilities in children with ADHD, suggesting age-appropriate MI abilities in this group and extensive individual variation in both TD and ADHD.

# Chapter 4: Investigating how within-task individual differences in MI impact the neural and behaviour correlates of VWM

## 4.1. Introduction

In an effort to examine how MI supports high-order cognition, research has investigated the relationship between MI and VWM. Evidence from neuroimaging research has implied shared visual representations between MI and VWM (Albers et al., 2013) and correlational studies with adults have shown MI is positively associated with VWM (Keogh & Pearson, 2011; 2014). However, studies, in the wider literature and in this thesis, are currently limited to comparing performance on MI and VWM measures. Data presented in **Chapter 2** and **Chapter 3** demonstrate that abilities in separable components of MI are distinct from VWM maintenance and VWM manipulation abilities in adults, typically developing children and children with ADHD. In addition, data-driven analysis in **Chapter 3** indicates individual differences in the profiles of both MI and VWM abilities in children with and without ADHD. Thus, a recurring interpretation throughout this thesis so far is that the recruitment of visual strategies in VWM is likely subject to extensive individual differences. In a similar vein, a recent review argued that to address competing theories of the neural underpinnings of VWM, i.e., a prominent role of high-level regions (e.g. Sreenivasan et al., 2014) vs. low-level regions (Serences, 2016), research should focus on understanding individual differences in the strategies recruited during VWM (Pearson & Keogh, 2019). Inferences regarding the extent to which visual strategies are recruited in VWM are currently speculative, given than research thus far has only compared performance on separate MI and VWM tasks. In order to establish how visual strategies are recruited in VWM, and thus derive conclusions based on the role of MI in VWM, methods are required to examine *how* MI is recruited *within* a VWM task. This investigation is vital if we are to delineate the types of strategies that are recruited in VWM and if we are to understand how MI might support memory and learning throughout development and in atypical groups. In the first known study of its kind, this chapter will adapt a classic VWM, orientation-discrimination task to investigate how individual differences in the precision of visual representations and the sensory experience of MI impacts neural correlates and behavioural outcomes of VWM. As this is the first known paradigm of its kind to investigate how MI modulates VWM capacity, the study involves a sample of typical adults in order to examine a well-established ERP component of VWM capacity; CDA. This ERP component has been shown to vary between children, young adults, and older adults (Sander et al., 2011). Therefore, in order to provide the groundwork to developing a theoretical framework of how the visual quality and capacity of MI modulates VWM, it is necessary to examine this relationship with an established ERP component to inform the design of future studies with a range of ages. Overall, this study directly examines the current speculations regarding how individual differences in MI underpin VWM abilities and extends the research presented thus far in this thesis comparing performance between MI and VWM tasks.

### 4.1.1. The relationship between MI and VWM

The investigation of how MI and VWM are related is relatively limited, yet the suggestion that they are similar functions is based on the parallels between the definitions of MI and VWM and the evidence for functional activation underpinning the two abilities (Lorenc et al., 2015; Miller & D'Esposito, 2005; Sreenivasan et al., 2014; Spagna et al., 2021). Both MI and VWM are conceptualised as involving the maintenance and manipulation of visual information (Baddeley, 2003; Baddeley & Andrade, 2000; Cowan, 2001; Logie, 1995; Kosslyn et al., 2006). Moreover, much like in the MI neuroimaging literature (reviewed in detail in **Chapter 1 Section 1.1.1**), there is evidence for a functional role of the frontal regions in VWM (Miller & D'Esposito, 2005; Sreenivasan et al., 2014) but there is also evidence that the visual regions play an important role in VWM (Serences, 2016). This has since raised the conceptual issue of whether MI and VWM are in fact the same processes but described in parallel literatures (Tong, 2013). On the other hand, it has been argued that conflicting findings regarding the importance of either frontal or visual regions in VWM is likely dependent on individual differences in the recruitment of visual strategies in VWM (Pearson & Keogh, 2019). It is therefore speculated that not all individuals approach *visual* memory tasks in the same way, however the direct investigation of how individuals use visual strategies in a VWM task is yet to be conducted.

Although there may be individual differences in strategies recruited, there is evidence for shared visual representations between MI and VWM. Findings have firstly shown that oriented gratings held in mind in VWM can be decoded using MVPA in visual areas V1-V4 (Harrison & Tong, 2009). This has then been extended to show that a classifier trained on early visual area activation in VWM trials reliably decoded activation in MI trials and vice versa (Albers et al., 2013). Based on this evidence, we might conclude that MI and VWM are therefore not distinct (Tong, 2013), however, behavioural evidence does not entirely align with this suggestion. Behavioural studies adopting a sensory strength measure of MI have implied that the recruitment of visual strategies in VWM is dependent on MI ability. Previous findings suggested that presenting visual noise during the delay period of a VWM task negatively impacts performance, which is taken to suggest it disrupts the visual information from being held in mind (Andrade et al., 2002). This interpretation is supported by the finding that MI is also disrupted when background luminance is modulated (Pearson et al., 2008). In turn, it has been shown that VWM performance was significantly poorer in the modulated background luminance condition but only in those that scored highly on the MI sensory strength measure. It was therefore interpreted that only “good imagers” recruit visual strategies in VWM (Keogh & Pearson, 2011; 2014). In support of this finding, a case study of an individual with Aphantasia found performance on a VWM task was intact (Jacobs 2018). The evidence thus far therefore implies the recruitment of MI in VWM may be dependent on variability in MI ability.

In addition, the findings reported in **Chapter 2** of this thesis that show the generation and maintenance of visual mental images are not significantly associated with VWM maintenance and manipulation supports the notion of individual differences in recruitment of visual strategies in VWM. A possible explanation for the seemingly contradictory findings of shared mechanisms between MI and VWM in the neuroimaging study (Albers et al., 2013) compared to dissociated components of MI and VWM found in this thesis so far could be the difference in methodological approaches. The paradigm presented in Albers et al. (2013) was a change detection task with either a VWM cue (i.e., in this trial, the participant is asked to view the stimulus and hold it in mind prior to detecting the change in the probe stimulus) or a MI cue (i.e., in this trial, the participant must imagine rotating the previously viewed stimulus prior to detecting the change). The fact that the participant is instructed to anticipate both MI and VWM trials might encourage the recruitment of visual strategies in the VWM trials, which would explain the evidence for shared visual representations. In contrast, MI and VWM are examined *separately* in **Chapter 2** and **Chapter 3** and the order of tasks was counterbalanced. In the image generation and image maintenance tasks, participants are instructed to generate an image (image generation) or hold an image (image maintenance) of the abstract shape in mind to determine if the probe shape looked different to the imaged shape. Comparatively, the instructions for the VWM task do not prime the participant to specifically recruit a visual strategy; they are instructed to click on the lily pads the frog appeared on either front to back (forward span/VWM maintenance) or back to front (backward span/VWM manipulation). The fact that participants were not primed to recruit a visual strategy in the VWM task likely led to individual differences in the types of strategies recruited in the VWM task. This is not the first study to suggest this. A recent study that examined the effects of training a visualisation strategy for a set of VWM tasks in adults found that in the control group (no strategies trained) only 4% reported visualisation (e.g., "I visualised the numbers") and no participants in the control group reported a self-generated imagery strategy (e.g., "I tried to associate each digit with some image in my mind.") (Forsberg et al., 2020). Instead, self-generated strategies included rehearsal ("I repeated the list of letters in my mind"), grouping ("I remembered the digits in groups") and other ("I made up a song…"). Taken together, this is therefore in line with the argument presented in Pearson and Keogh's (2019) review that individuals likely recruit different strategies in their approach to VWM tasks and that those with strong MI might be more inclined to recruit visual strategies in VWM, which could in turn aid performance (Keogh & Pearson, 2011; 2014). Examining the extent to which individual differences in MI impact VWM would further elucidate the role of visual strategies/MI in supporting memory. Research thus far has been restricted to comparisons between absolute performance on MI measures and on VWM measures, therefore, to fully elucidate how MI supports VWM, it is necessary to investigate how within-task individual differences in the precision of visual representations and the sensory experience of MI impacts VWM performance.

### 4.1.2. Measuring the precision and number of items held in mind during VWM

A primary line of enquiry in examining the brain-behaviour mechanisms of VWM is the investigation of how content-specific information is maintained in VWM as quantified by EEG correlates. In a seminal study, Vogel and Machizawa (2004) identified neural correlates of VWM maintenance in the occipital and parietal electrodes. Visual memory echoes the organisation of the visual system in that memory traces are more prominent in the hemisphere contralateral to the presentation of the stimuli (i.e., the memory trace of a stimulus presented on the left side of the screen is more prominent in the right visual cortex) (Gratton, 1998). Vogel and Machizawa (2004) presented a four-item memory array to each hemifield to capitalise on the contralateral functioning of the visual areas and establish hemispheric activity corresponding to lateralised visual arrays. Specifically, they found a sustained negative waveform in the hemisphere contralateral to the presented stimuli during the delay period (900ms) following the memory array (100ms). Next, the number of items in the memory array was varied between trials and a difference wave was computed (contralateral activity minus ipsilateral activity), which was termed contralateral delay activity (CDA). Notably, it was found that as set size increases, the CDA becomes more negative (and therefore increases) up to an asymptote of 4 items. This crucial finding that CDA indexes VWM capacity has since been replicated extensively (see Luria et al., 2016 for review). CDA is typically observed in the posterior electrodes, namely P5/6, P7/8, PO3/4, PO7/8, and O1/2. It is calculated as the mean amplitude of difference wave (contralateral activity minus ipsilateral activity) during the delay period from around 400ms following sample onset to the end (~900 to 1400ms) of the delay period (McCollough et al., 2007; Vogel & Machizawa, 2004). In sum, it is possible to track the up to 4 visual items held in VWM with millisecond precision using EEG.

Findings have since been extended to show that the precision at which representations are held in VWM is mapped by fluctuations in CDA at smaller set sizes. It should be noted here that the term precision used in **Chapters 2** and **3** is not conceptually different from the use of the term precision in Machizawa et al. (2012) described below. In this thesis thus far, the term precision has been used to refer to the visual precision at which visual information is held in mind as indexed by the detection of changes in a probe shape. In Machizawa et al. (2012), the term precision is used to characterise the visual precision at which visual information is held in mind as indexed by the detection of either fine (15º) orientation or coarse (45º) orientation changes in a probe array. Namely, an orientation-discrimination paradigm was adopted whereby precision of the visual representation of items could be characterised by an interaction between degrees of target orientation (15º fine orientation change vs. 45º coarse orientation change) and set size (2 and 4 items). Here it was shown that at a set size of 2 items, but not 4 items, CDA was larger in fine orientation change compared to coarse orientation change trials (Machizawa et al., 2012). Thus, it is apparent that at smaller set sizes, CDA reflects variation in the precision of information held in mind during the VWM delay, suggesting that CDA not only maps the number of items in mind but is modulated by precision of the representation of items before reaching an asymptote at around 4 items. Notably, what might be described in the VWM literature as precision of representations, would ultimately be described as the visual vividness or quality of mental images in MI literature. We might therefore assume that at smaller set sizes, neural correlates of precision, i.e., CDA, reflects the visual quality of visual images held in mind during VWM and otherwise CDA reflects the number of visual items held in mind. However, this has not been quantified alongside the reported subjective, sensory experience of MI. Thus, knowledge of how the visual quality and quantity of items held in a representation can impact VWM performance and modulate CDA amplitudes is incomplete.

### 4.1.4. Considering the functional role of frontal regions in MI and VWM

As outlined above, there is contention over the functional role of frontal vs. visual regions in VWM. Comparatively, recent research in MI has demonstrated the importance of selective activation in the frontal regions in the construction of mental images (reviewed in detail in **Chapter 1 Section 1.2.1.2**) with specific reference to reversed directionality of activation in MI compared to visual perception (Breedlove et al., 2020; Dentico et al., 2014; Dijkstra et al., 2020). The investigation of how individual differences in MI impact VWM would therefore be incomplete without also considering the role of early frontal ERP components. Conclusions regarding the contribution of specific regions as measured by event-related potentials (ERPs) using EEG should be interpreted cautiously given the restricted spatial resolution of EEG. However, analysis of ERPs from frontal compared to posterior electrodes can give an indication of the relative contribution of such areas (e.g., Murray et al., 2011). Research has identified an electrophysiological marker of preparatory attention in VWM following the cue at the very beginning of VWM trials. Namely, the anterior directing attention negativity (ADAN) component in the frontal electrodes is argued to index the initiation of voluntary shifts in attention following cue onset and in preparation of the memory array (Nobre et al., 2000). This is argued to reflect prioritisation of sensory information in VWM (Myers et al., 2017). ADAN is measured in the frontal electrodes, namely FC3/4 and C3/4 (Jongen et al., 2007). Investigation into the role of preparatory attention in visual short-term memory (VSTM) has shown that ADAN is associated with a greater number of items held in VSTM, which is taken to suggest evidence for facilitation of encoding of items in VSTM (Murray et al., 2011). While the study of the role of preparatory attention in MI is limited, one study has compared conditions of MI and VSTM and found more ADAN-like effects in the MI condition compared to the VSTM condition (Gosling & Astle, 2013). It is therefore important to consider this research in the context of evidence that demonstrates selective activation of the frontal regions precedes activation in the visual areas during MI (e.g., Dijkstra et al., 2020). Namely, it might be anticipated that greater ADAN following cue onset reflects the initiation of a vivid MI strategy within a VWM task, however the relationship between ADAN and MI within a VWM task is yet to be examined directly. This would provide further evidence for a functional role of the frontal regions in MI and would disentangle the role of frontal regions in VWM with respect to individual differences in MI.

### 4.1.4. Measuring MI within a VWM task

The research presented in this thesis thus far has adopted a separable-component model of MI, however, the most common approach to MI research is to measure the subjective sensory experience of MI using ratings. This is not surprising given that MI is an inherently private and variable sensory experience. In the quest to establish evidence to suggest that visually depictive representations are recruited during MI, research has examined the relationship between the subjective, sensory experience of MI and selective activation of visual areas. As outlined in **Chapter 1, Section 1.2.1.1.**, the sensory experience of MI is most often measured using the Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973; Marks, 1995). The VVIQ includes a series of 16 statements whereby participants are required to visualise the content of the statement and rate how vivid their mental image is on a scale of 1-5 ranging from 1 = “perfectly clear and as vivid as normal vision” to 5 = “no image at all, you only ‘know’ you’re thinking of an object”. There are various categories of statements including natural scenes, e.g., “Visualise a rising sun. Consider carefully the picture that comes before your mind’s eye”, and familiar faces. Neuroimaging studies have demonstrated a positive relationship between scores on the VVIQ and selective activation of the early visual areas during MI (Cui et al., 2007; Lee et al., 2012). This therefore suggests that the activation of visual areas, and thus the recruitment of visual representations in MI, is directly related to the subjective experience of the vividness of MI.

The findings above have since been extended in paradigms adopting trial-by-trial vividness ratings. The first known study to examine trial-by-trial vividness ratings (1 = almost no imagery, 2 = some weak imagery, 3 = moderate imagery, 4 = strong imagery almost like perception) found a positive association between the MI sensory strength score, i.e., the extent to which perception in a binocular rivalry paradigm is biased following an imagery period (Pearson et al., 2008), and trial-by-trial subjective vividness ratings (Pearson et al., 2011). This was interpreted to suggest that individuals have good insight into their MI. More recently, it has been shown that the overlap between regions activated during MI and visual perception is positively associated with trial-by-trial subjective vividness (1 = not vivid at all to 4 = very vivid; Dijkstra et al., 2017b). Taken together, the findings imply that the subjective sensory experience of the vividness of MI maps onto precision of visual representations. However this has not been directly assessed with respect to VWM. Moreover, while evidence in the VWM literature suggests that the number of items held in mind during the delay period in VWM can be quantified by CDA, the extent to which this reflects the subject sensory experience of MI is yet to be addressed. Therefore, adapting a VWM paradigm to include trial-by-trial subjective vividness ratings and quantity ratings (number of items in mind) would present a novel opportunity to address the current gap in the literature in understanding how individual differences in MI impact VWM.

### 4.1.4. The current study

The current study was designed to directly examine how MI is recruited in a VWM task in the form of two clear aims. The first aim is to characterise the visual precision and capacity of items held in VWM maintenance by extending previous findings to not only consider how CDA and behavioural performance vary dependent on precision and capacity conditions, but also how subject ratings vary dependent on precision and capacity in VWM. The second aim is to establish the metacognitive link between the subjective sensory experience of MI and behavioural and neural correlates of VWM: CDA and ADAN. To address these two aims, a novel VWM paradigm was adapted to include trial-by-trial subjective vividness and quantity ratings. This paradigm was derived to establish whether subjective vividness ratings, i.e., the sensory experience of the visual quality of mental images, map onto neural correlates of precision in VWM (CDA) and to establish whether subjective quantity ratings, i.e., the sensory experience of the number of visual items in mind, maps onto the neural correlates of capacity (CDA) in VWM maintenance. Finally, the relationship between individual differences in vividness and quantity ratings and ADAN will be examined to establish to metacognitive link between MI and preparatory attention within VWM. Together, the findings will demonstrate how MI is recruited in VWM.

For clarity, the term vividness is recruited in line with previous definitions in that participants will rate how vivid their representation is (e.g., Pearson et al., 2011; Marks, 1973). The term quantity refers to the number of items held in mind. It might also be questioned why MI is tested as a singular concept in this chapter following the evidence from **Chapters 2** and **3** suggesting MI is a multi-faceted construct. For the purpose of the investigation in this chapter, MI is treated as the visual quality (vividness and quantity) of representations in order to establish how MI/visual strategies are recruited in VWM.

Prior to outlining the hypotheses, it is important to note here that the first paradigm piloted in this study (detailed in the methods) aimed to introduce a vividness and/or quantity rating to the original orientation-discrimination paradigm introduced in Machizawa et al. (2012). However, following piloting, an instruction condition was included in the main experiment, which is additional the original Machizawa et al. (2012) paradigm. A recent study with a similar paradigm to Machizawa et al. (2012) paired a capacity-focused instruction with coarse precision (45° orientation-change) trial blocks and a quality-focused instruction with fine precision (15° orientation-change) trial blocks. Briefly and preceding a more detailed explanation in the methods section, in the current study, instruction and ratings were manipulated so that a capacity-focused instruction was paired with vividness ratings in one block and quantity ratings in another block and the quality-focused instruction was paired vividness ratings in one block and quantity ratings in another. In the original study conducted by Machizawa et al. (2012), fine and coarse orientation discriminations (precision trials herein) were *expected* based on colour, therefore participants were aware following the sample onset of the level of precision required for the trial. In the main experiment in the current study, fine and coarse trials were randomised within each block, therefore instruction was manipulated so that participants knew to *expect* to maintain either a representation of high visual precision (quality-focused) or to maintain as many items as possible (capacity-focused). The introduction of an instruction condition therefore allows for investigation of how individuals flexibly recruit representations of varying precision and capacity, as well as determining whether the neural correlates of precision and capacity in VWM map onto the subjective sensory experience of MI. This is the first known study to manipulate instruction; hypotheses for the instruction condition are detailed below. For clarity, quality is used to refer to the quality-focused condition within the factor of instruction, and precision is used to refer to the factor of precision (orientation discrimination), however visual quality and visual precision are not conceptually different.

With reference to the first aim, behavioural outcomes are firstly expected to replicate previous findings. Proportion correct is measured with respect to correctly identified orientation-changes. It is hypothesised that previous findings will be replicated in that there will be higher proportion correct in smaller set sizes compared to larger set sizes as well as in coarse compared to fine precision trials (Machizawa et al., 2012). As outlined above, blocks were separated by instruction and rating (block 1: quality-focused and vividness rating, block 2: capacity-focused and vividness rating, block 3: quality-focused and quantity rating, block 4: capacity-focused and vividness ratings) and fine and coarse trials were randomised within each block. An interaction between instruction and precision is expected in proportion correct in that greater proportion correct is expected in coarse precision compared to fine precision trials in quality-focused trials only. Instruction is also expected to modulate subjective ratings in that greater vividness ratings are expected in quality-focused blocks compared to capacity-focused blocks and greater quantity ratings are expected in capacity-focused blocks compared to quality-focused blocks. In turn, if the assumption that precision and capacity in VWM equates to the visual precision and number of visual items held in MI holds true, instruction is expected to modulate subjective ratings as described above.

Measuring EEG during the behavioural VWM task allows for the unique opportunity to directly measure the visual precision and capacity of items held in mind during the delay period (via CDA) and to measure preparatory attention following cue onset (ADAN). CDA is therefore expected to reveal further nuances of the visual precision and capacity of representations, over and above behavioural outcomes. As CDA is measured during the delay period and before the behavioural response, if CDA is modulated by instruction this will demonstrate that individuals can flexibly control the precision and capacity of their visual representations based on instruction. If this is the case, differences in CDA are expected between quality-focused trials compared to capacity-focused trials but not between fine and coarse precision trials, this is because participants *expect* and can prepare for either quality- or capacity-focused responses but fine and coarse trials are not cued therefore they cannot prepare for this. In sum, this will extend previous findings by examining how instruction modulates VWM consumption as indexed by CDA amplitude. Finally, it is also expected that the established CDA set size effect will be replicated here in that CDA will increase as a function of set size up to 4 items.

Next, individual differences in the metacognitive link between MI and behavioural and neural correlates of VWM will be investigated. Evidence for significantly greater accuracy in trials rated as high vividness compared to low vividness is expected and significantly greater accuracy in non-divergent ratings (e.g., rated 2 items in mind when required to remember 2 items) compared to divergent quantity ratings (e.g., rated 2 items in mind when required to remember 4 items) is also predicted. With regard to neural correlates, it is predicted that CDA amplitudes will be significantly larger in high vividness trials compared to low vividness trials and this effect is likely to be greater in quality-focused trials at smaller set sizes. It is also predicted that CDA amplitudes will be significantly larger in larger set sizes in trials with non-divergent quantity ratings compared to trials with divergent ratings. Together, this would support the assumption that individuals have good metacognitive insight into their visual representations and extend this to insight into MI during VWM. Moreover, it will demonstrate that CDA not only maps the visual precision and/or capacity of representations, but also the subjective sensory experience of MI within VWM. To examine how preparatory attention is linked to the subjective sensory experience of MI, ADAN amplitudes will be examined. It is firstly predicted that larger ADAN amplitudes will be observed in high vividness trials compared to low vividness trials, which would support the hypothesis for a role of preparatory attention in MI with respect to the visual precision of MI. Moreover, it is predicted that larger ADAN amplitudes will be observed in non-divergent quantity rating trials compared to divergent quantity rating trials, which would further support the hypothesis that preparatory attention plays a role in the MI with respect to the number of visual items held in mind. Together, this would therefore provide a novel method for quantifying the role of MI in VWM.

## 4.2. Methods

### 4.2.1. Participants

Participants were recruited from the SONA database at Birkbeck, University of London and the surrounding community. All participants gave written informed consent and were either paid £25 to participate or received course credit. Participants had normal or corrected-to-normal vision and each participant completed the Ishihara 38 Plates CVD Test (https://www.color-blindness.com/ishihara-38-plates-cvd-test/) to check for red-green colour deficiencies and were required to score “none” to participate. A total of 7 individuals participated in the pilot (age: *M* = 26.40, *SD* = 5.94, 2 female) and a total of 23 individuals were recruited for the final experiment. Prior to artefact rejection, two participants were excluded due to incomplete datasets due to technical errors and three more participants were excluded as they did not respond in any of the trial-by-trial ratings and thus did not produce any behavioural ratings data. One more participant was excluded following artefact rejection due to there being less than 75% of the total trials remaining. Therefore, a total of 17 participants are included in the reported results for the main experiment (age: *M* = 26.00, *SD* = 4.39, 10 female).

### 4.2.2. Materials and procedure: pilot

Tasks were comprised and presented with MATLAB v2016b (The MathWorks Inc, Natick, MA) and Cogent 2000. Initially, the first pilot involved the simple modification of adding vividness and quantity ratings to the original paradigm presented in Machizawa et al. (2012). However, including a subjective rating prior to the behavioural orientation-discrimination response was too difficult for participants. Alterations were therefore made to the paradigm to reduce the influence of task difficulty. For completeness, the pilot procedure is reported below followed by pilot results and discussion before reporting the main experiment procedure. EEG was recorded during the piloting; however, it is not reported here.

4.2.2.1. Pilot procedure

The VWM paradigm was initially designed to replicate the paradigm presented in Machizawa et al. (2012) with three modifications: the addition of either a vividness rating or quantity rating within the trial (collectively referred to as a subjective rating herein), the addition of a 1 item set size to act as a baseline, and the randomisation of the precision (fine and coarse trials) condition within the block as in Machizawa et al. (2020). The trial sequence including timing information is presented in Figure 4.1.

Participants were instructed to memorise an array of bars, hold the bars in mind and subsequently determine whether the highlighted bar in the probe array had been rotated clockwise or counter-clockwise. A fixation point was presented in the centre of the screen throughout the trial and participants were required to maintain their gaze at the fixation point. First, participants were cued to memorise either the bars presented to the left or right side of the screen. Second, the sample display was presented which consisted of 2, 4 or 8 bars (1, 2 or 4 bars presented to each hemifield, respectively). The colour of the bars in the sample display indicated whether the participant should anticipate a fine or a coarse orientation change. Red bars indicated a fine precision trial (15° orientation change) and green bars indicated a coarse precision trial (45° orientation change). Participants were instructed to maintain fixation at the central fixation point and hold the bars in mind as accurately as possible during the subsequent delay. Third, the rating screen appeared which displayed the question “vividness?” in half the blocks and “quantity?” in the other half of the blocks. The order of blocks was counter-balanced, and participants were informed of the rating type at the start of the block. In vividness rating blocks, participants were required to rate the vividness of the image in mind from a scale of 1 to 4 in line with previous paradigms (Pearson et al., 2011): 1 = almost no image, 2 = weak image, 3 = moderate image, 4 = strong image/almost like perception. The options were displayed underneath the “vividness?” question on the screen. In quantity rating blocks, participants were required to rate the number of items they held in mind. If participants had to remember 1 item, the “quantity?” question was displayed with the options of 0 or 1 beneath. Pressing 0 would indicate no visual memory of the items, 1 would indicate participants felt they accurately remembered the 1 item. If participants were required to remember 2 items, 0, 1 or 2 were displayed as options and if participants were required to remember 4 items, they had the option to respond with either 0, 1, 2, 3 or 4. Finally, participants were presented with the probe array which was the same as the sample array except the highlighted bar had been rotated either clockwise or counter-clockwise and participants used the left and right arrow to answer whether the bar had been rotated clockwise or counter-clockwise (see Figure 4.1 for schematic presentation of trial sequence).

Participants completed 10 experimental blocks for the pilot (three participants completed 8 blocks and one participant completed 6 blocks due to technical errors). Each block consisted of 48 trials totalling to 480 trials over 10 blocks, thus the number of trials per set size (1, 2 and 4) was 160 and number of trials per level of precision (fine and coarse) was 240. The order of conditions (set size, precision, direction of orientation) was randomised in each block and each trial type occurred the same number of times in each block. There were 12 orientations which were randomly assigned to each bar and ranged in 15° intervals from 5° to 170°, therefore vertical, horizontal, and diagonal orientations were not included. Items were presented with a minimum item-to-item distance of 2° visual angle, with the maximum visual angle to the closest edge from fixation point set at 8° and the minimum set at 4°. Participants sat 65cm from the screen. Practise consisted of one block.

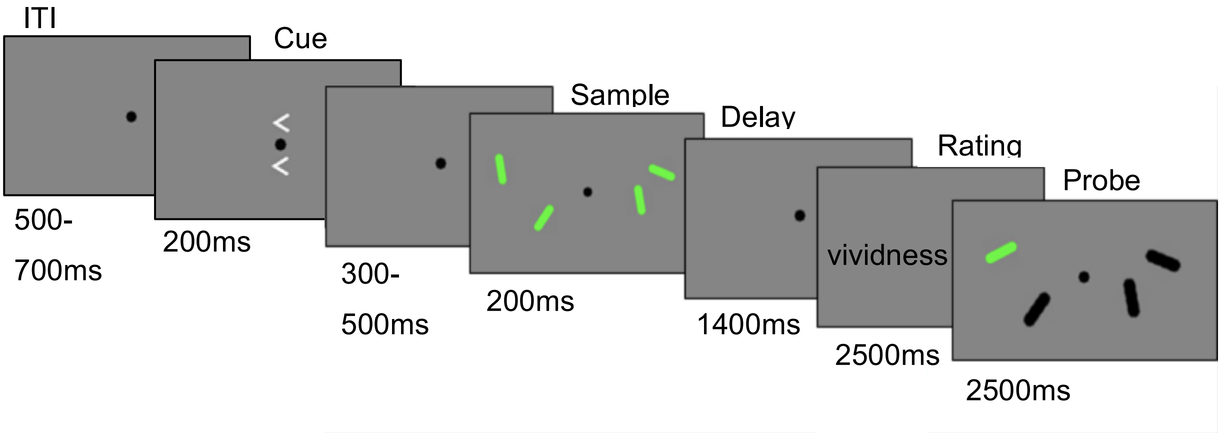



*Figure 4.1:* Trial sequence. Inter-trial interval ranged between 500-700ms. For each trial, an arrow cue was presented for 200ms to indicate which side of the screen should be attended to. This was followed by a 300-500ms interval before the sample array was presented for 200ms. The sample array consisted of 1, 2 or 4 bars on each side of the screen (set size 2 pictured) and either red (fine precision condition) or green (coarse precision condition pictured) bars. This was following by a 1400ms delay period whereby participants had to hold the image in mind. After the delay, participants provided either a vividness or quantity rating (vividness pictured). Subsequently, a probe array until the participant responded (or 2500ms) whereby all stimuli except the target stimulus was presented in black. Participants were required to judge whether the target was rotated clockwise (pictured) or counter-clockwise compared to the sample array

4.2.2.2. Pilot results and discussion

Overall mean proportion correct was .59 (*SD* = .04). Mean and standard deviations of proportion correct for each set size and precision condition found in the pilot and those reported in Machizawa et al. (2012) are outlined in Table 4.1. As can be observed in the table, proportion correct in the pilot was much lower than in Machizawa et al. (2012). Together, this suggests that the subjective rating rendered the task too difficult therefore a second experimental paradigm was designed to reduce the influence of task difficulty.

Table 4.1.

*Pilot 1: Means and standard deviations of proportion correct in each condition from the pilot and Machizawa et al. (2012)*

| | Fine | Coarse | Fine (Machizawa et al., 2012) | Coarse (Machizawa et al., 2012) |
|---|---|---|---|---|
| Set size 1 | .68 (.07) | .59 (.05) | N/A | N/A |
| Set size 2 | .59 (.07) | .60 (.09) | .72 (.10) | .76 (.11) |
| Set size 4 | .53 (.03) | .55 (.04) | .66 (.09) | .61 (.07) |

*Note. Machizawa et al. (2012) did not include set size 1, hence N/A is reported.*

### 4.2.3. Materials and procedure: main experiment

For the main experiment, several changes were made to ensure participants could complete the task and render sufficient trials for analysis. First, blocks were differentiated by instruction and rating whereby vividness and quantity ratings were presented in either a quality-focused or a capacity-focused block, respectively. This procedure was adapted from Machizawa et al. (2020) where a quality-focused instruction was applied for fine precision trials and a capacity-focused instruction for coarse precision trials. The effect of the two different instructions on VWM performance and CDA amplitude has not been examined previously and is therefore novel to this paradigm. In the current study, fine and coarse trials were randomised within each block and were not cued by colour. Rather than attributing colour to fine and coarse precision trials, colour was attributed to instruction: green bars presented in capacity-focused instruction blocks and red bars presented in quality-focused instruction blocks. In the quality-focused blocks, participants were instructed to hold a precise image in mind and in the capacity-focused block, participants were instructed to focus on holding as many items in mind as they could (i.e., they should try and hold all 4 items in mind in the 4-item condition). Participants were warned at the start of the experiment that they may have to discriminate either fine or coarse orientation-discriminations and there would be no cue for this.

A total of four blocks of 96 trials (384 total trials) were presented with two breaks within each block to reduce fatigue and boredom. Block 1 was quality-focused with vividness ratings, block 2 was capacity-focused with vividness ratings, block 3 was quality-focused with quantity ratings and block 4 was capacity-focused with quantity ratings. The order of blocks was counterbalanced per participant. Set size (1 item, 2 items, 4 items), precision (fine, coarse) and attended side (left, right) were randomised within each block resulting in eight trials per condition. A recent study conducted stimulations to estimate how many participants and how many trials are required for different levels of power in CDA analyses (Ngiam et al., 2021). It was suggested that 30-50 trials were required per condition to detect the presence of CDA and up to 400 trials per condition with 25 participants could be needed to detect differences between set size conditions in CDA with 80% power. The task with 384 trials already takes just under an hour to complete, therefore adding more trials would distort the quality of the data. Moreover, robust CDA effects have been established in previous studies with ~20 subjects and ~80 trials per condition (Machizawa et al., 2012, 2020). Therefore, it is possible to be confident in the number of trials for at least main effects and 2-way interactions reported in the ANOVA with all conditions in the results. For example, subject to artefact rejection, for the 3-level factor of set size, there would be maximum 128 trials per level and for 2 level factors such as instruction, there would be maximum 192 trials per level. However, 3-, 4- and 5-way interactions should be interpreted with caution due to limited trial numbers. To familiarise participants with the task, they completed a quality-focused block and capacity-focused block (with either vividness or quantity ratings, counterbalanced) with 24 trials per block as a practise. The practise blocks were repeated if participants scored < 65% percentage correct.

A confidence rating was included at the end of each block where participants were asked to rate their confidence in their behavioural performance of that block. This was to check for the relationship between confidence rating of orientation discrimination performance and subjective ratings of vividness and quantity. While the subjective rating is purposefully placed before the probe array in the trial sequence to reduce the confound of confidence, a weak correlation between metacognitive ratings and confidence is expected. To test this, participants were presented with a blank grey screen at the end of the block with "confidence?" in the centre and they were required to answer according to a standard 5-point Likert scale: 1 = not confident at all, 2 = slightly confident, 3 = somewhat confident, 4 = fairly confident, 5 = completely confident (4 confidence trials in total).

The second change made to the paradigm following piloting was to the rating cue. Previous evidence has shown that VWM performance is negatively impacted in trials where another visual stimulus that requires a response is presented within the delay period (Bae & Luck, 2019). Therefore, instead of displaying a visual cue of "vividness?" or "quantity?", the screen remained the same as the delay period with only the fixation point presented and a tone cue was used (similar to previous MI paradigms in neuroimaging studies, e.g., Stokes et al., 2009) to indicate to participants that they now had to rate their representation. The tone was generated in Cogent 2000 and comprised a 250Hz sine wave lasting 200ms, which was played from a speaker placed behind the participant's chair. The trial sequence is depicted schematically in Figure 4.2.

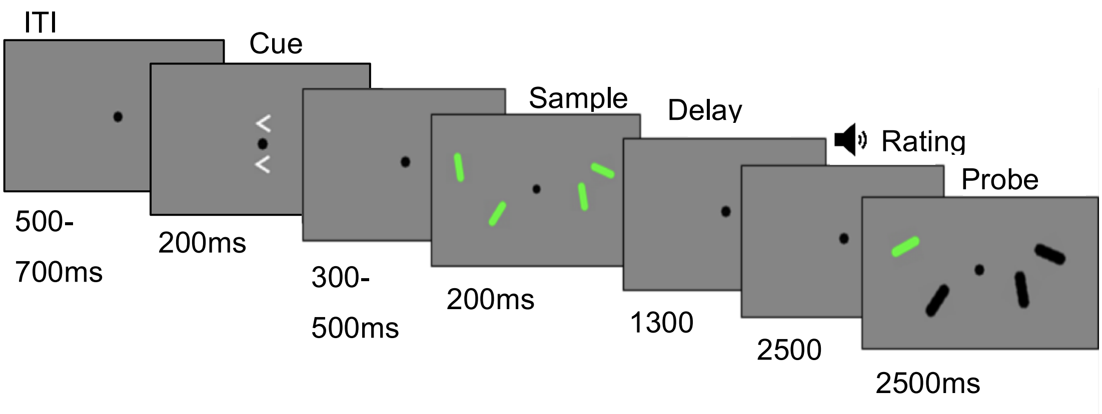


*Figure 4.2:* Trial sequence of experimental paradigm. The main experiment paradigm sequence is exactly as outlined in Figure 4.1, except that the rating cue is now presented as a tone. Example of left cue, set size of 2, rating, and coarse, clockwise orientation-discrimination

Finally, distances between items were altered. When the task is difficult, participants tend to saccade to the array of items they are required to attend to, however, the nature of the CDA analyses require participants to remain fixated at the central fixation points (McCollough et al., 2007; Vogel & Machizawa, 2004). Therefore, to reduce the propensity to saccade to the stimulus, the visual angle of item presentation was moved closer to the fixation point and the item size was increased to aid encoding. The minimum distance from the edge of the fixation point was reduced from 4° to 2.5° and the maximum distance from the edge of the fixation point was reduced from 8° to 6.5° with a minimum item-to-item distance increase from 2° to 2.5°. This was to account for the increase in item size from 1.5° visual angle to 2°. The size of cues and fixation point remained the same.

#### 4.2.3.1. EEG recording

EEG data was continuously recorded offline at 1,000Hz sampling rate using a fitted cap (EASYCAP) with 64 Ag-AgCl passive electrodes according to the international 10-20 system using a BrainVision amplifier. The cap included two horizontal EOG channels mounted in the cap and a vertical EOG channel was placed directly underneath the right eye to monitor blinks and saccades. Electrical impedance was kept below 5 kΩ. During the recording, FCz acted as the reference electrode and AFz as the ground electrode.

#### 4.2.3.2. Measuring horizontal eye movements

In most EEG studies, horizontal EOG channels are commonly placed directly next to the eyes (outer sides of canthi), however in the EASYCAP system, the horizontal EOG channels are built into the cap and are therefore placed horizontally in line with eyes but on the side of the head. As horizontal saccades distort assumed positions of targets on the retina fundamental for bilateral ERP components such as CDA and ADAN, it was important to determine the magnitude of the amplitude of saccades at the horizontal electrode sites so that the saccades can be properly detected and removed prior to analysis. Conventionally, when EOG electrodes are positioned around canthi, a horizontal EOG amplitude at 25μV is set as a threshold to detect and reject horizontal eye movements in bilateral target displays in periphery (e.g., McCollough et al., 2007; Vogel & Machizawa, 2004). As the horizontal EOG channels in the EASYCAP are positioned further away from the eyes compared to previous EEG methods in CDA research, it was necessary to set appropriate horizontal EOG rejection criteria in the current study paradigm. To do this, five of the participants who participated in the main experiment completed a 5-minute saccade task (depicted in Figure 4.3) prior to participating in the main experiment. This was then used to calculate the threshold for saccade detection in the horizontal EOG (HEOG) channels. In the saccade task, participants were required to focus on the fixation square in the centre of the screen. Next, a red square appeared either 2, 4, 6, 8 or 10 degrees to the left or right of the central fixation square. Participants were instructed to keep their eyes on the central fixation square until they saw a red square, at which point they were required to saccade to the red square and look back to the central black square and wait for the next trial. The task involved a total of 200 trials (50 per condition) over 5 blocks.

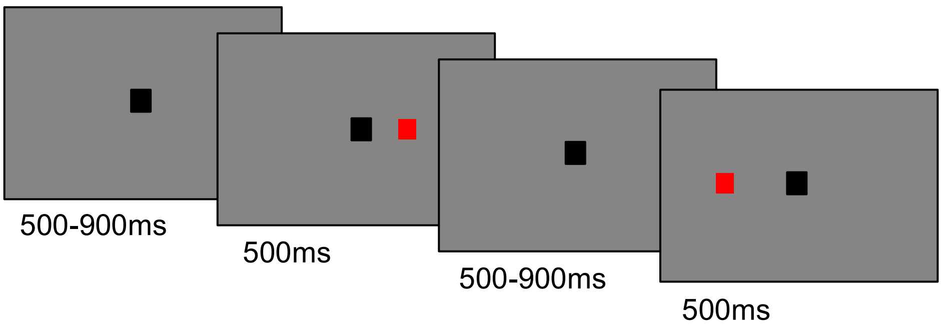
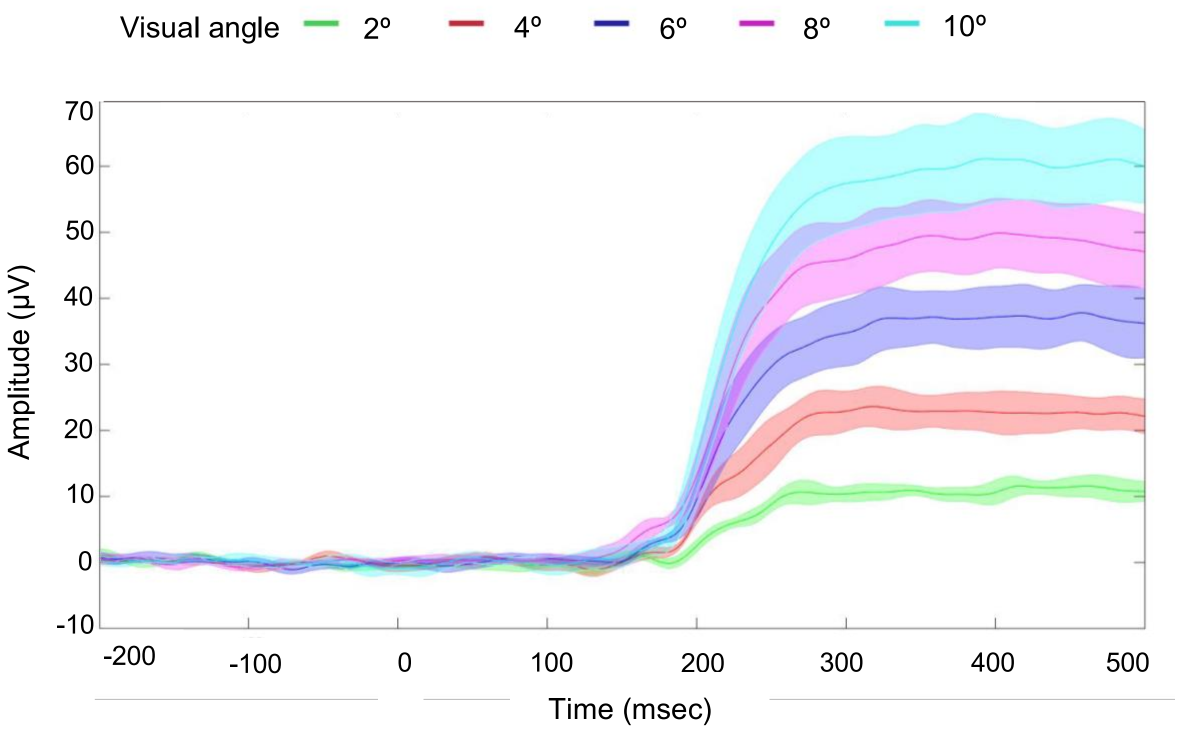



*Figure 4.3:* Sample trial sequence for two trials in the saccade task. Central fixation presented for random delay between 500-900ms, followed by the presentation of the red square that participants had to saccade to

A bipolar HEOG channel was derived (right horizontal EOG channel subtracted from left horizontal EOG channel) to observe the magnitude of left and right saccades, respectively. Mean amplitudes between 300-500ms following cue-onset were calculated for each visual angle (2º: $M$ = 10.88, $SD$ = 2.43; 4º: $M$ = 22.82, $SD$ = 6.07; 6º: $M$ = 36.91, $SD$ = 10.04; 8º: $M$ = 48.65, $SD$ = 12.21; 10º: $M$ = 59.94, $SD$ = 14.09). The waveforms are depicted in Figure 4.6., which suggest larger amplitudes are observed when participants saccade at




greater visual angles. This was supported by ANOVA of amplitude with a within-subject factor of visual angle (2º, 4º, 6º, 8º, 10º), which revealed a significant main effect of visual angle ($F(4,16) = 73.82$, $p < .001$, $\eta_p^2 = .94$). Post hoc comparisons revealed no overlap across visual angles ($p$s < .001). Based on the mean amplitudes, a simple formula can be applied to estimate the degree of horizontal eye movement: y = x / 6; where x = bipolar HEOG channel amplitude and y = degrees the eyes moved. The stimuli in the main experiment were presented between 2.5º -6.5º visual angle and the formula indicates that 2º saccades would be characterised as a mean bipolar HEOG channel amplitude of +/- 12µV and 3º would be characterised as mean bipolar HEOG channel amplitude of +/- 18µV. To avoid overcorrection of data, 18µV was chosen as the final value to detect saccades in the main experiment trials.

*Figure 4.4:* Mean and SE of amplitude in the bipolar HEOG channel for each visual angle (2º, 4º, 6º, 8º, 10º) where the stimulus is presented at 0 msec

4.2.3.3. Artefact detection

After the recording, data was pre-processed in MATLAB (2016b) using the MATLAB toolbox EEGLAB (version 2019.1.; Delorme & Makeig, 2004). Data were filtered offline with an 8th-order Butterworth bandpass filter at 0.05-30Hz and resampled at 500Hz. For CDA analyses, data were epoched to -200 to 1400ms around the sample array onset and baseline corrected (-200-0ms). Blinks during the sample array onset (0-200ms) were first detected using a moving window peak to peak detection algorithm with a window size of 200ms, a step of 10ms and a threshold of 50µV, trials with blinks during the sample array onset were then rejected (*M* = 23, *SD* = 18, *range* = 4 to 61).

Next, an algorithm to detect square waves in the bipolar HEOG channel was applied with the threshold criteria set to +/-18µV. While this was effective in detecting saccades, the algorithm also detected +/-18µV square waves that were too quick to be saccades (i.e., 50ms) (mean number of trials detected = 86, *SD* = 64, *range* = 7 to 200). Therefore, the trials flagged by the algorithm were checked by eye to determine whether the square waves detected were in fact saccades, i.e., the square wave spanned ~200ms (mean number of trials detected = 32, *SD* = 34, *range* = 7 to 118). As can be seen from the range, if all trials with saccades were removed, this would result in more participants being excluded due to insufficient data. Research has shown that applying independent component analysis (ICA) to remove saccade and blink components does not distort data for CDA analyses and is therefore an efficient method to retain data (Drisdelle et al., 2017). ICA was therefore conducted using the SOBI algorithm in EEGLAB and components were observed using ICLabel (Pion-Tonachini et al., 2019). Saccade and blink components were detected with the aid of ICLabel, which labels components according to the pattern of activity (e.g., eye component, muscle components etc.). An average of 2 (*SD* = 1, *range* = 1 to 5) components that were deemed either blink or saccade components were removed.

The blink and saccade algorithms were re-applied to the ICA corrected data and any remaining trials with saccades exceeding the 18µV threshold and blinks exceeding the 50μV threshold were rejected (*M* = 4, *SD* = 2, *range* = 0 to 8). Finally, extreme values of +/-75μV and abnormal trends of linear drift (50μV, R = .80) were detected and rejected and the average number of remaining trials for the CDA analyses following all artefact rejection was 335 (*SD* = 30, *range* = 278 to 364). The data was then re-referenced to mastoid electrodes, in line with previous literature conducting CDA analyses (Machizawa et al., 2012). Channels rejected due to noise by EEGLAB automated criteria were interpolated (*M* = 1, *SD* = 1, *range* = 0 to 2). Finally, the average CDA component was obtained from P5/6, P7/8, PO3/4, PO7/8, and O1/2 channels and was computed from 400-1400ms after sample onset. The number of trials remaining following artefact rejection was similar across all conditions (see Tables A.3.1-A.3.3 in the **Appendix**).

For ADAN analyses, filtered data were epoched to -200-500ms post arrow cue onset and baseline corrected (-200-0ms). Blinks and saccades were detected using the same algorithms applied in CDA artefact rejection (pre-ICA mean number of trials with blinks detected = 54, *SD* = 36, *range* = 8 to 115; mean number of trials with saccades detected = 38, *SD* = 30, *range* = 0 to 90). ICA was conducted as in the CDA analyses and an average of 1 (*SD* = 1, *range* = 1 to 3) components were removed. Following ICA, extreme values and abnormal trends were detected as above and trials containing artefacts were rejected (*M* = 6 *SD* = 7, *range* = 0 to 44). The blink and saccade algorithms were reapplied to the ICA-corrected data and any remaining trials with blinks and saccades were removed. The average number of remaining trials for the ADAN analyses following all artefact rejection was 336 (*SD* = 32, *range* = 274 to 377). The data was then re-referenced to the mastoid electrodes in line with previous pre-processing procedures in research examining ADAN (Murray et al., 2011). Given the short time window, channel noise was low, and one channel was rejected for one participant and interpolated. Finally, the average ADAN component was obtained from FC3/4 and C3/4 and was computed from 350-500ms (Jongen et al., 2007; Murray et al., 2011). The number of trials remaining following artefact rejection was similar across all conditions (see Table A.3.4. in the **Appendix**).

#### 4.2.3.4. Analysis strategy

Tests of normality revealed some variables were not normally distributed, however parametric analyses were applied given that ANOVA is robust to violations of assumptions of normality (Blanca et al., 2017). All within-subject post hoc comparisons are reported with Bonferroni corrections. Where assumptions of sphericity were violated, Greenhouse-Geisser estimates are reported. Where variables were not normally distributed, Spearman's correlations are reported and where variables were normally distributed, Pearson's correlations are reported. RTs for subjective ratings and behavioural responses to the probe array that were less than 250ms or equal to 2500ms (no response) were not included in analyses.

## 4.3. Results

### 4.3.1. Characterising the visual precision and capacity of VWM maintenance as indexed by proportion correct, subjective MI ratings and CDA

#### 4.3.1.1. Proportion correct

Overall accuracy (as measured by proportion correct) was sufficient for the main experiment (*M* = .71, *SD* = .09) and descriptive statistics of proportion correct for all conditions are reported in Tables 4.2.A-C.

Table 4.2.A.

*Means and standard deviations of proportion correct for each condition for 1 item trials*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean | .84 | .86 | .86 | .89 | .79 | .79 | .79 | .80 | .85 | .84 | .87 | .87 | .78 | .85 | .83 | .79 |
| (SD) | (.18) | (.14) | (.13) | (.13) | (.18) | (.16) | (.11) | (.18) | (.16) | (.18) | (.15) | (.12) | (.18) | (.15) | (.19) | (.18) |

*Note.* Q = Quantity, V = Vividness

Table 4.2.B.

*Means and standard deviations of proportion correct for each condition for 2 item trials*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean | .73 | .69 | .72 | .81 | .74 | .71 | .66 | .66 | .72 | .65 | .78 | .72 | .64 | .72 | .72 | .76 |
| (SD) | (.20) | (.14) | (.22) | (.19) | (.18) | (.18) | (.16) | (.18) | (.21) | (.16) | (.23) | (.19) | (.25) | (.24) | (.13) | (.14) |

*Note.* Q = Quantity, V = Vividness

Table 4.2.C.

*Means and standard deviations of proportion correct for each condition for 4 item trials*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean | .49 | .56 | .64 | .65 | .55 | .57 | .61 | .57 | .69 | .63 | .64 | .62 | .63 | .56 | .53 | .62 |
| (SD) | (.26) | (.19) | (.27) | (.18) | (.24) | (.22) | (.23) | (.21) | (.23) | (.21) | (.21) | (.23) | (.16) | (.21) | (.18) | (.22) |

*Note.* Q = Quantity, V = Vividness

A repeated measures ANOVA was conducted with proportion correct as the dependent variable and within-subject factors of set size (1 item, 2 items, 4 items), precision (fine, coarse), instruction (capacity-focused, quality-focused), rating (vividness, quantity), and attended side (left, right). Firstly, there was a significant main effect of set size ($F(2,32) = 82.59$, $p < .001$, $\eta_p^2 = .84$), Bonferroni corrected post hoc comparisons revealed a significant decrease in proportion correct between all comparisons (all *ps* < .001). Also in line with previous findings, there was a significant main effect of precision ($F(1,16) = 10.31$, $p = .005$, $\eta_p^2 = .39$), such that there was greater proportion correct in coarse precision (45° orientation-change) trials compared to fine precision (15° orientation-change) trials. There was no main effect of instruction ($F(1,16) = 2.58$, $p = .128$, $\eta_p^2 = .39$), rating ($F < 1$) or attended side ($F(1,16) = 2.58$, $p = .128$, $\eta_p^2 = .39$). There was a trend for an interaction between precision and instruction ($F(1,16) = 3.31$, $p = .088$, $\eta_p^2 = .17$). Given that the 2-way interactions are arguably powered in terms of trials and there were a priori hypotheses regarding the interaction between precision and instruction, the interaction was explored and reported with caution. Post hoc t-tests revealed no difference between fine and coarse trials in the capacity-focused blocks ($t(16) = 1.16$, $p = .263$, $d = .28$) but there was significantly higher proportion correct in coarse trials compared to fine trials in the quality-focused blocks ($t(16) = 3.57$, $p = .003$, $d = .87$), in line with hypotheses (see Figure 4.5).

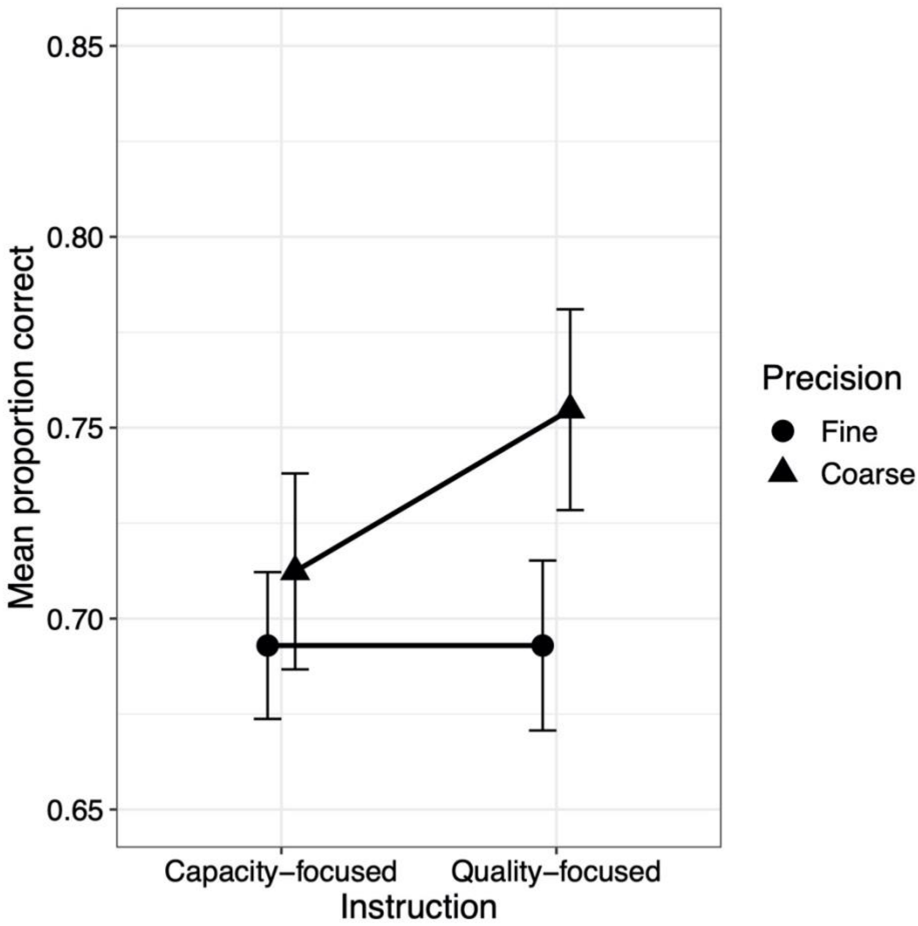



*Figure 4.5:* Means and SE of proportion correct for precision (fine, coarse) and instruction (capacity-focused, quality-focused)

A trend for a 3-way interaction between attended side, precision and rating was found ($F(1,16) = 3.97$, $p = .064$, $\eta_p^2 = .19$). Follow up ANOVAs of each precision condition (fine and coarse, respectively) revealed a significant interaction between attended side and rating in the coarse precision trials only ($F(1,16) = 5.33$, $p = .035$, $\eta_p^2 = .25$) (fine precision trials: attended side x rating interaction: $F < 1$). This was further followed up using t tests to determine the effect of attended side for each rating type. This revealed significantly greater proportion correct in right ($M = .76$, $SD = .12$) attended trials compared to left attended ($M = .71$ , $SD = .13$) trials in the quantity ratings condition ($t(16) = 2.22$, $p = .042$, $d = .54$), alongside no significant difference between right attended ($M = .72$, $SD = .11$) and left attended ($M = .74$, $SD = .09$) trials in the vividness ratings condition ($t(16) = 1.34$, $p = .199$, $d = .33$). Thus, again the source of the interaction relates to differential effects of attended side, with stronger performance in right attended trials, this time differentiated by rating type.

There was also a significant 3-way interaction between attended side, set size and instruction ($F(2,32) = 4.31$, $p = .022$, $\eta_p^2 = .21$). Follow up ANOVAs for each set size were conducted to explore this interaction. There was a significant interaction between attended side and instruction only in the 4-item condition ($F(1,16) = 7.22$, $p = .016$, $\eta_p^2 = .31$) (1-item condition attended side x instruction interaction: $F < 1$; 2-item condition attended side x instruction interaction: $F(1,16) = 1.88$, $p = .189$, $\eta_p^2 = .11$). Follow up t tests revealed an effect of attended side; significantly greater proportion correct in the right attended trials compared to the left attend trials for the capacity-focused condition ($t(16) = 3.01$, $p = .03$, $d = .36$), but not in the quality-focused condition ($t(16) = .58$, $p = 1.00$, $d = .07$). Thus, the three-way interaction is best explained as differential effects of attended side, for each instruction condition in the 4-item condition only.

Finally, there was a trend for a 4-way interaction between attended side, set size, instruction and rating ($F(3,32) = 2.98$, $p = .065$, $\eta_p^2 = .16$). This revealed that the three-way interaction above between attended side, instruction and set size was driven by quantity rating blocks only ($F(1,32) = 6.50$, $p = .004$, $\eta_p^2 = .29$) (vividness rating block: attended side x instruction x set size interaction: $F < 1$). Follow up ANOVAs for each set size demonstrated the same pattern as for the 3-way interaction above. That is, a significant interaction between attended side and instruction in the 4-item condition only ($F(1,16) = 10.10$, $p = .006$, $\eta_p^2 = .02$) (1-item attended side x instruction interaction: $F < 1$; 2-item attended side x instruction interaction: $F(1,16) = 3.46$, $p = .081$, $\eta_p^2 = .18$). As with the 3-way interaction above, this was best explained via t tests of attended side to show significantly greater proportion correct in right attend compared to left attended trials in the capacity-focused condition ($t(16) = 2.99$, $p = .009$, $d = .73$), but not between right and left attend trials in the quality-focused condition ($t(16) = 1.21$, $p = .245$, $d = .29$) (see Figure 4.6).

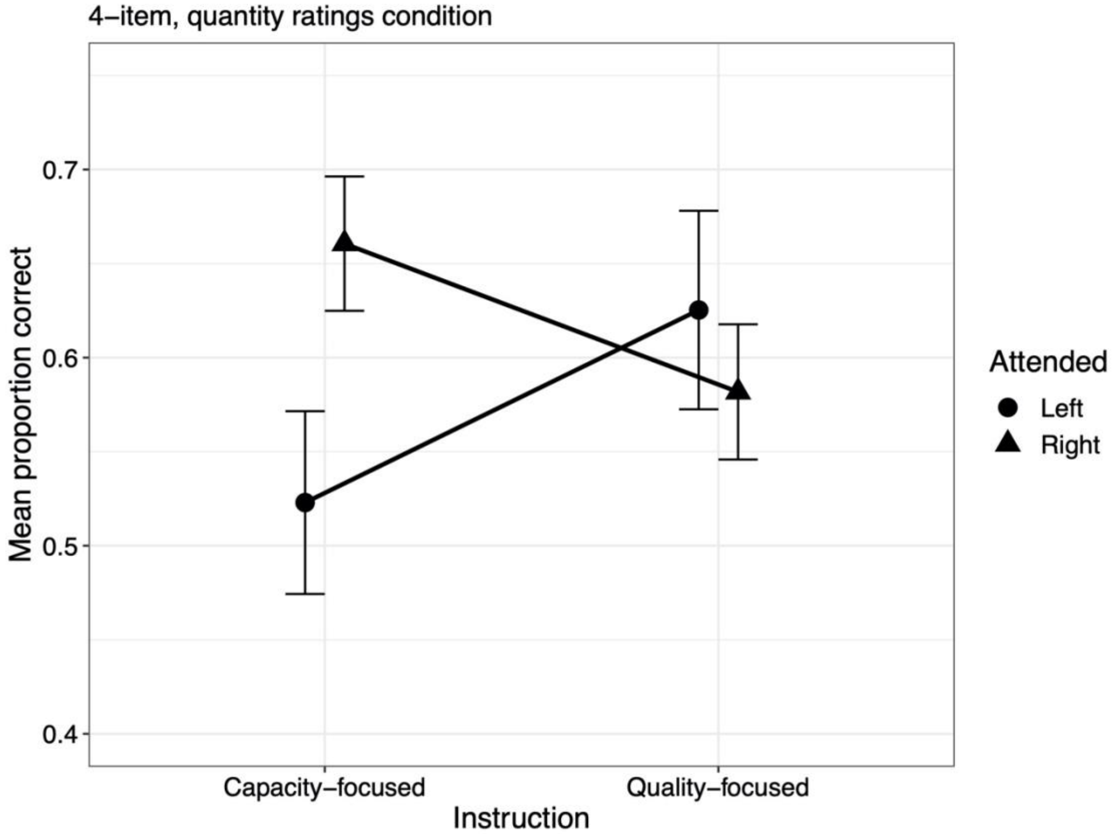



*Figure 4.6:* Mean and SE of proportion correct for 4-item quantity trials only with attended side (left, right) and instruction (capacity-focused, quality-focused)

4.3.1.2. Metacognitive ratings

Next, separate ANOVAs were conducted on vividness ratings and quantity ratings, respectively. The within-subject factors were set size (1 item, 2 items, 4 items), precision (fine, coarse), instruction (capacity-focused, quality-focused) and attended side (left, right). Descriptive statistics of vividness ratings per condition are presented in Tables 4.3.A-C.

Table 4.3.A.

*Means and standard deviations of vividness ratings per condition for 1 item trials*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | Fine | | Coarse | | Fine | |
| Instruction | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused |
| Mean (SD) | 2.82 (1.05) | 2.82 (.95) | 2.90 (1.06) | 2.73 (.99) | 2.90 (1.06) | 2.81 (1.11) | 2.98 (1.00) | 2.84 (1.02) |

Table 4.3.B.

*Means and standard deviations of vividness ratings per condition for 2 item trials*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | Fine | | Coarse | | Fine | |
| Instruction | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused |
| Mean (SD) | 2.96 (.76) | 2.73 (.93) | 2.82 (.79) | 2.76 (.73) | 2.82 (.73) | 2.62 (.90) | 2.79 (.61) | 2.65 (.88) |

Table 4.3.C.

*Means and standard deviations of vividness ratings per condition for 4 item trials*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | Fine | | Coarse | | Fine | |
| Instruction | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused |
| Mean (SD) | 2.37 (.64) | 2.24 (.93) | 2.37 (.77) | 2.28 (.84) | 2.44 (.89) | 2.29 (.92) | 2.51 (.79) | 2.24 (.91) |

The vividness ratings ANOVA revealed a significant main effect of set size ($F(2,32) = 3.58$, $p = .04$, $\eta_p^2 = .18$), where post hoc comparisons showed marginally significantly higher vividness ratings when participants were required to remember 1 item compared to when they remembered 4 items ($p = .055$) (all other $ps > .05$) (Figure 4.7). There was no main effect of precision ($F < 1$) and no main effect of attended side ($F < 1$). There was a trend for a main effect of instruction ($F(1,16) = 3.71$, $p = .072$, $\eta_p^2 = .18$), where the means implied higher vividness ratings for the capacity-focused blocks compared to quality-focused blocks. Finally, there were no significant interactions (all $F < 1.06$, n.s.).

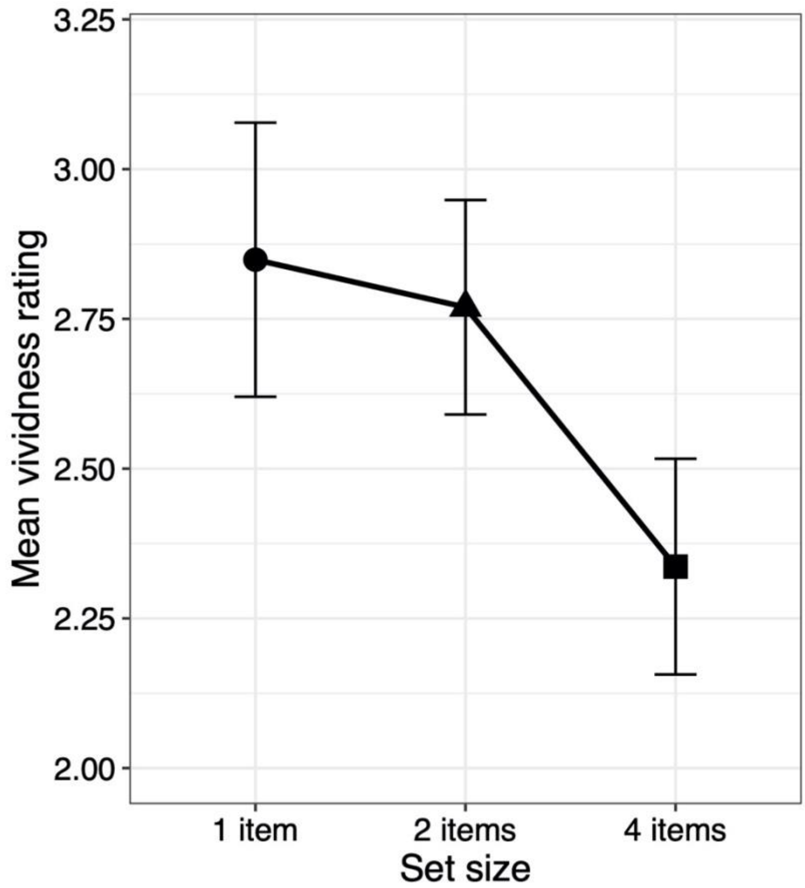



*Figure 4.7:* Mean and SE of vividness ratings for each set size (1 item, 2 items, 4 items)

An equivalent ANOVA was conducted on quantity ratings with the same within-subject factors as the vividness ratings ANOVA. Descriptive statistics for quantity ratings for each condition are reported in Tables 4.4.A-C.

Table 4.4.A.

*Means and standard deviations of quantity ratings per condition for 1 item trials*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | Fine | | Coarse | | Fine | |
| Instruction | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused |
| Mean (SD) | 1.03 (.26) | 1.02 (.24) | 1.08 (.34) | .99 (.32) | 1.04 (.27) | 1.00 (.22) | .99 (.35) | .103 (.26) |

Table 4.4.B.

*Means and standard deviations of quantity ratings per condition for 2 item trials*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | Fine | | Coarse | | Fine | |
| Instruction | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused |
| Mean (SD) | 1.88 (.27) | 1.79 (.42) | 1.95 (.32) | 1.83 (.36) | .183 (.32) | 1.73 (.33) | 1.85 (.29) | 1.81 (.35) |

Table 4.4.C.

*Means and standard deviations of quantity ratings per condition for 4 item trials*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | Fine | | Coarse | | Fine | |
| Instruction | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused | Capacity-focused | Quality-focused |
| Mean (SD) | 2.46 (1.06) | 2.46 (1.15) | 2.52 (1.03) | 2.42 (1.13) | 2.50 (.09) | 2.40 (1.03) | 2.55 (.95) | 2.41 (1.12) |

There was a significant main effect of set size ($F(2,32) = 32.04$, $p < .001$, $\eta_p^2 = .67$), where quantity ratings significantly increased with each increase in number of items (2 items > 1 item: $p < .001$, 4 items > 2 items: $p = .004$, 4 items < 1 item: $p < .001$; Figure 4.8). There were no main effects of precision, instruction or attended side (*Fs* < 1) and there were no significant interactions (all *F*s < 1, n.s.).

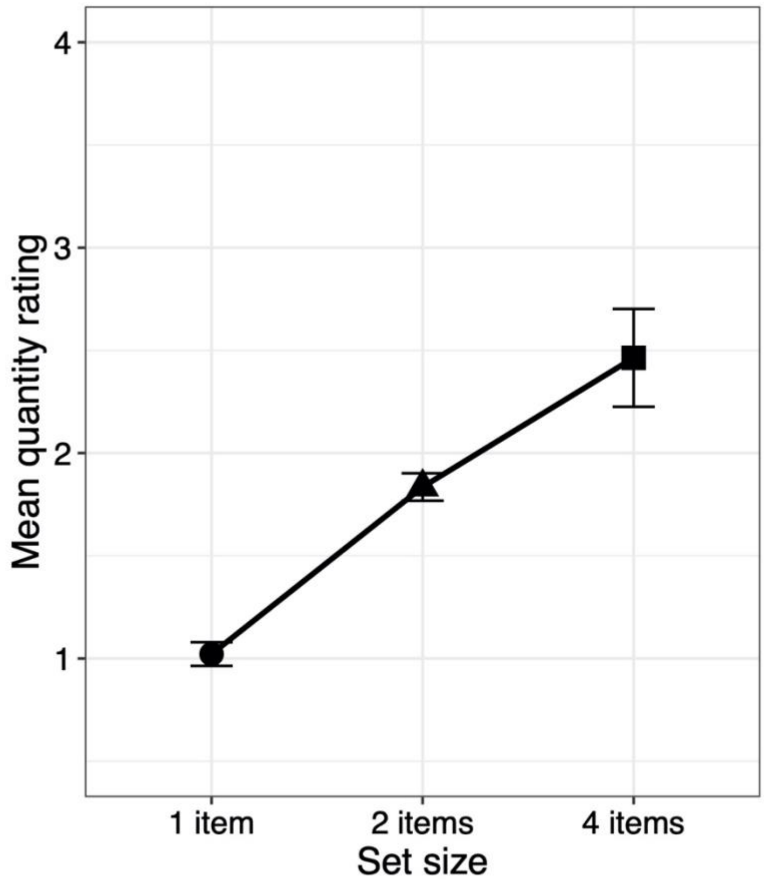



*Figure 4.8:* Mean and SE of quantity ratings for each set size (1 item, 2 items, 4 items)

4.3.1.3. Contralateral delay activity (CDA)

To examine how CDA was modulated by condition, an ANOVA was conducted on grand-averaged CDA and within-subject factors of set size (1 item, 2 items, 4 items), precision (fine, coarse), instruction (capacity-focused, quality-focused), rating (vividness, quantity) and attended side (left, right). Descriptive statistics for all conditions are presented in Tables 4.5.A-C.

Table 4.5.A.

*Means and standard deviations of CDA for each condition for 1 item trials*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean | -1.07 | -1.13 | -0.75 | -1.40 | -1.12 | -1.39 | -1.05 | -0.89 | -0.98 | -0.64 | -0.34 | -1.22 | -1.02 | -0.58 | -0.82 | -1.38 |
| (SD) | (1.47) | (1.79) | (1.83) | (1.68) | (1.57) | (1.77) | (1.76) | (1.29) | (1.61) | (2.02) | (1.77) | (1.66) | (1.73) | (1.85) | (2.03) | (1.45) |

Table 4.5.B.

*Means and standard deviations of CDA for each condition for 2 item trials*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean | -1.48 | -1.19 | -1.29 | -1.25 | -1.82 | -1.15 | -1.46 | -1.27 | -1.35 | -1.29 | -1.94 | -1.29 | -0.73 | -1.27 | -1.23 | -0.98 |
| (SD) | (2.49) | (1.73) | (1.60) | (1.46) | (1.79) | (1.39) | (1.96) | (1.31) | (1.58) | (1.57) | (2.05) | (1.12) | (1.22) | (1.95) | (1.56) | (1.44) |

Table 4.5.C.

*Means and standard deviations of CDA for each condition for 4 item trials*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean | -2.14 | -1.55 | -1.85 | -1.34 | -2.13 | -1.97 | -1.57 | -1.76 | -1.26 | -2.00 | -0.91 | -1.53 | -0.73 | -1.21 | -1.49 | -1.84 |
| (SD) | (2.30) | (2.11) | (3.18) | (2.02) | (2.02) | (2.21) | (2.34) | (2.00) | (1.54) | (1.92) | (1.71) | (1.39) | (1.51) | (1.79) | (2.03) | (2.03 |

There was a significant main effect of set size ($F$(2,32) = 14.06, $p$ < .001, $\eta_p^2$ = .47). Post hoc comparisons revealed significantly greater CDA between 1 item ($M$ = -.99, $SD$ = .60) and 2 items ($M$ = -1.31, $SD$ = .59) ($p$ = .02) as well as 1 item and 4 items ($M$ = -1.58, $SD$ = .85) ($p$ < .001), and a trend for significantly greater CDA between 2 items and 4 items ($p$ = .07). There was no main effect of precision, instruction, rating or attended side (*Fs* < 1). There was a trend for an interaction between set size and rating ($F$(2,32) = 2.65, $p$ = .086, $\eta_p^2$ = .14). Follow up ANOVAs revealed a main effect of set size for both vividness ($F$(2,32) = 7.35, $p$ = .002, $\eta_p^2$ = .32) and quantity ratings ($F$(2,32) = 12.29, $p$ < .001, $\eta_p^2$ = .43). Pairwise comparisons in the vividness ratings ANOVA showed a significant increase in CDA amplitude for the largest set size of 4 items, compared to set sizes of 1 and 2 items (*p*s < .03), but not between the 1-item and 2-item trials ($p$ = 1.00). In the quantity ratings ANOVA, pairwise comparisons revealed that the significant increase in CDA amplitude was between set size 1, compared to set sizes of 2 and 4 (*p*s < .005), and not between the 2-item and 4-item trials ($p$ = 1.00) (see Figure 4.9).

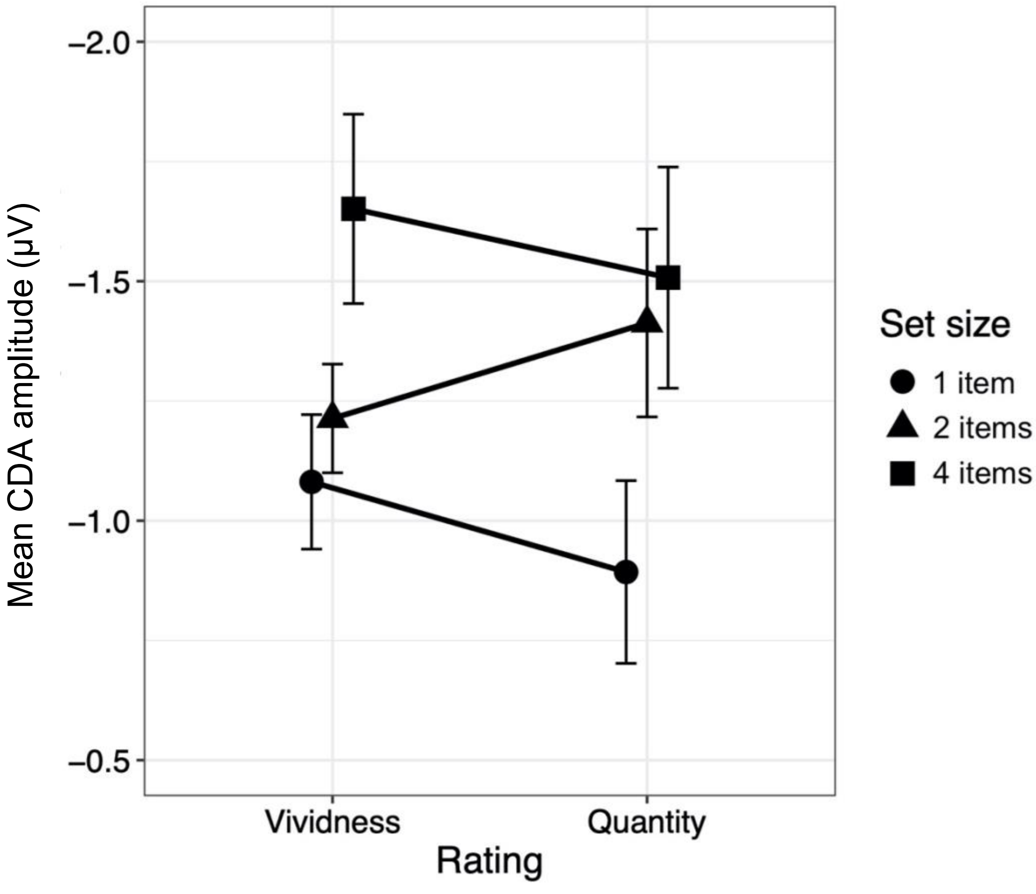



*Figure 4.9:* Mean and SE CDA amplitude for each rating type (vividness, quantity) and set size (1 item, 2 items, 4 items). Y axis reversed for reference

There was a significant 3-way interaction between precision, instruction and attended side ($F$(1,16) = 6.01, $p$ = .026, $\eta_p^2$ = .27) and a significant 4-way interaction between instruction, attended side, set size and rating ($F$(2,16) = 4.06, $p$ = .027, $\eta_p^2$ = .20). Follow up ANOVAs on quality-focused and capacity-focused blocks, respectively, were conducted to explore the significant 3-way interaction. There was a significant interaction between attended side and precision in the capacity-focused condition only ($F$(1,16) = 5.85, $p$ = .028, $\eta_p^2$ = .27) (quality-focused condition attended side x precision interaction: $F$ < 1). T tests of the effect of precision for each attended side revealed significantly greater CDA in coarse trials compared to fine trials in the right attend condition ($t$(16) = 2.74, $p$ = .015, $d$ = .66) but there was no difference between fine and coarse in the left attend condition ($t$(16) = 1.23, $p$ = .238, $d$ = .29).

With regard to the 4-way interaction, there was a significant interaction between attended side, set size and rating in the capacity-focused trials only ($F$(2,32) = 3.35, $p$ = .048, $\eta_p^2$ = .17) (quality-focused condition attended side x set size x rating interaction: $F$(1,32) = 1.41, $p$ = .259, $\eta_p^2$ = .08). Follow up ANOVAs for each set size for capacity-focused trials revealed a significant interaction between rating and attended size in the 2-item condition only ($F$(1,16) = 4.50, $p$ = .05, $\eta_p^2$ = .08) (1-item condition rating x attended side interaction: $F$(1,16) = 2.39, $p$ = .141, $\eta_p^2$ = .14; 4-item condition rating x attended side interaction: $F$(1,16) = 2.93, $p$ = .107, $\eta_p^2$ = .16). While the means point towards greater CDA amplitude in left ($M$ = -1.65, $SD$ = 1.93) compared right ($M$ = -1.04, $SD$ = .04) attend trials in the quantity ratings, this was not significant ($t$(16) = .97, $p$ = .345, $d$ = .24). There was also no significant difference between left attend ($M$ = -1.16, $SD$ = 1.35) and right attend trials ($M$ = -1.28, $SD$ = 1.59) in the vividness ratings condition ($t$(16) = .183, $p$ = .857, $d$ = .04). Although there was a 4-way significant interaction, the lack of differences found between individual conditions is likely dependent on low power due to the limited number of trials in each condition. There were no other significant interactions ($Fs$ < 1). Grand averaged ipsilateral, contralateral and CDA waveforms for each set size are presented in Figure 4.10 and grand-averaged CDA waveforms per set size for each block are presented in Figure 4.11.

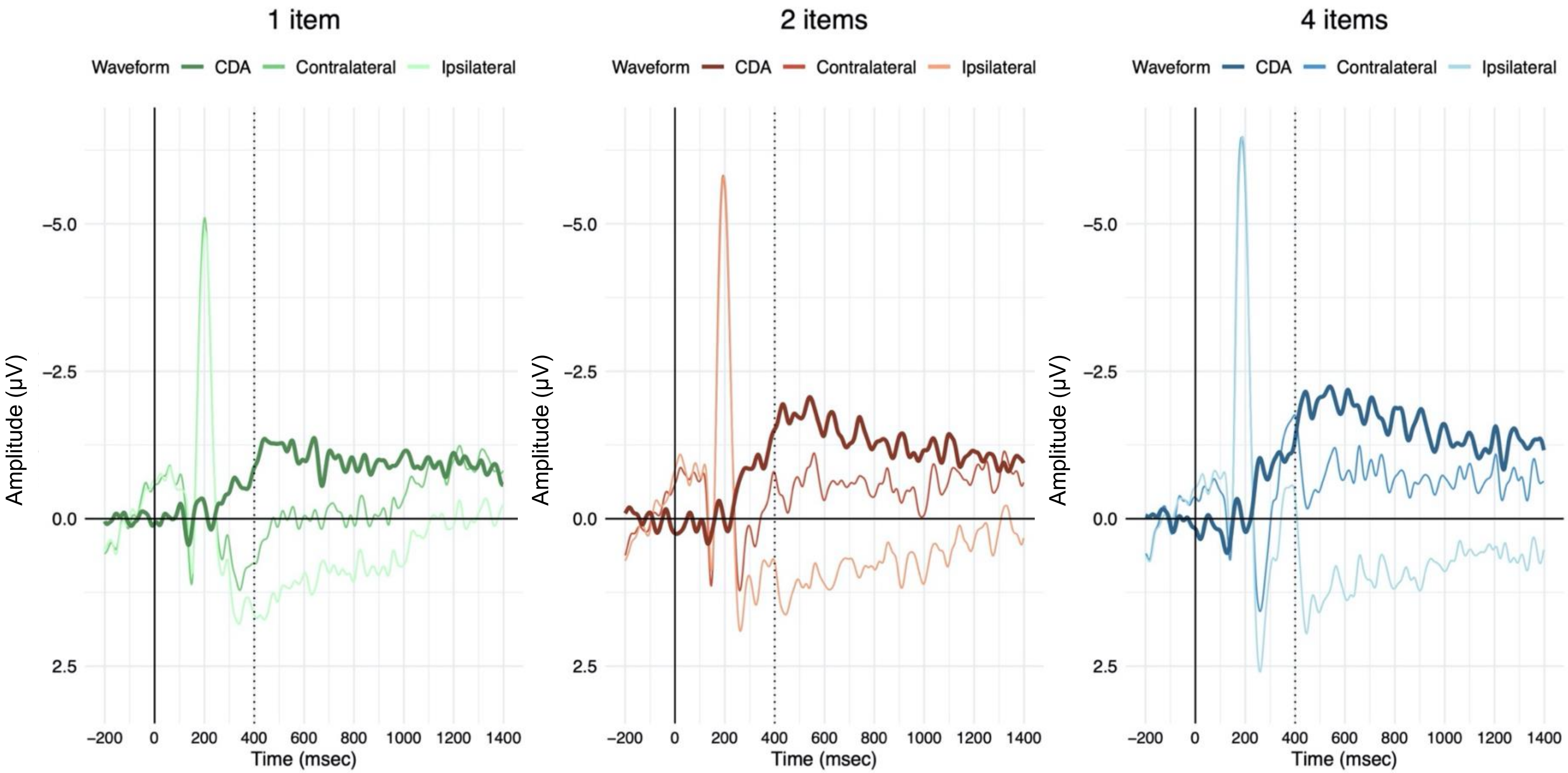



*Figure 4.10:* Grand-averaged waveforms for the 1 item trials (left), 2 items (centre) and 4 items (right). Sample onset is at 0-200msec and vertical dotted line at 400ms added for reference (CDA amplitude calculated as mean amplitude between 400ms and 1400ms after sample onset)

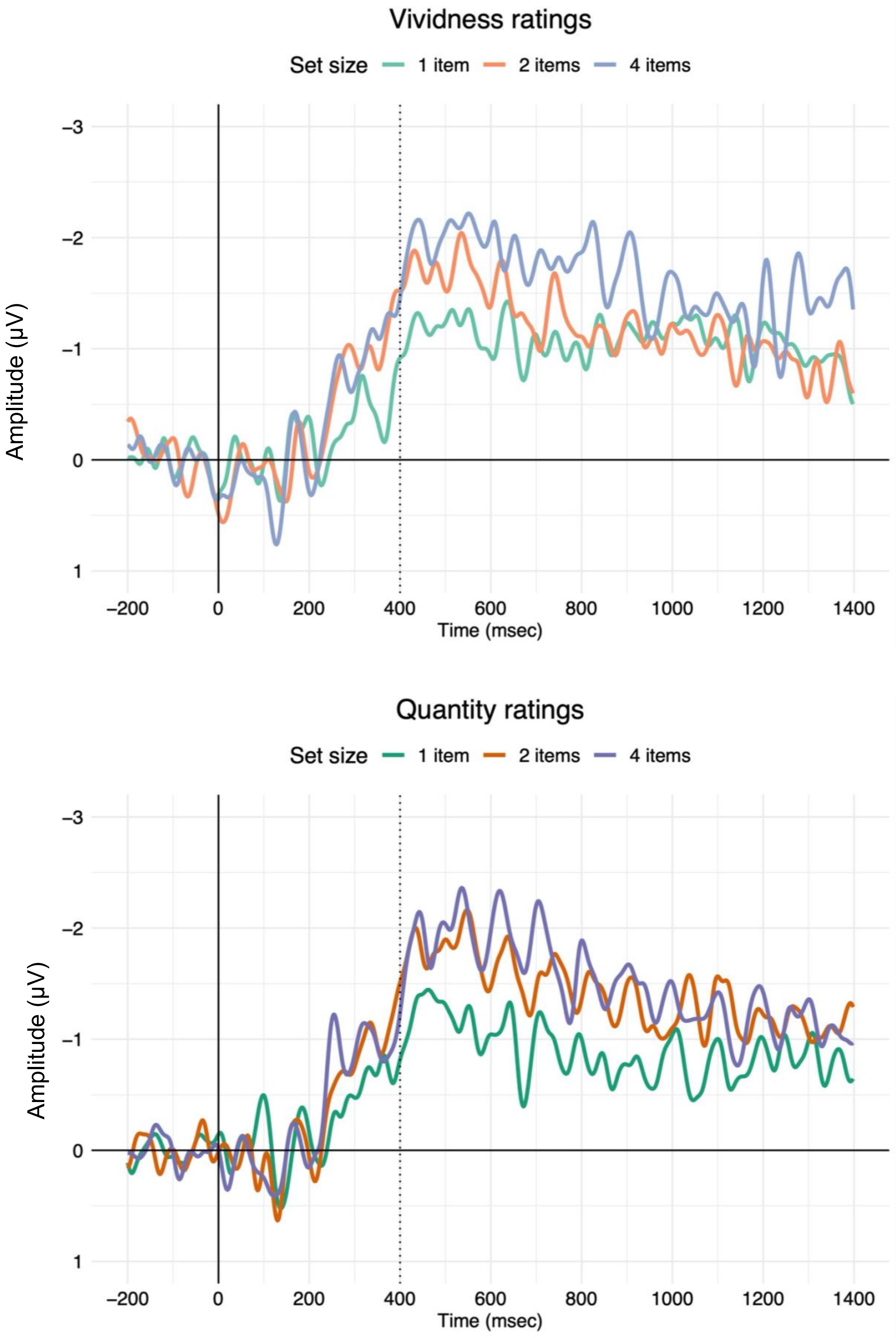



*Figure 4.11:* Grand-averaged waveforms of CDA for each set size (1 item, 2 items, 4 items) for vividness rating block (top) and quantity rating block (bottom)

### 4.3.2. Individual differences in the metacognitive link between MI and behavioural and neural correlates of VWM

#### 4.3.2.1. Relationship between subjective MI ratings and confidence ratings

A Pearson's correlation was conducted between mean confidence ratings for vividness rating blocks (*M* = 3.53, *SD* = .62) and mean vividness ratings (*M* = 1.58, *SD* = .26). This revealed a strong positive correlation between confidence ratings and vividness ratings ($r = .508$, $p = .037$), which suggests the higher participants rated vividness, the greater the confidence participants had in their VWM accuracy. The equivalent Pearson's correlation was conducted between mean confidence ratings for quantity blocks (*M* = 3.47, *SD* = .91) and mean quantity divergence score (*M* = .58, *SD* = .40). A mean divergence score was calculated based on divergent and non-divergent responses. Non-divergent responses were scored 0 and were trials where the participant rated that they had all items in the array clearly in mind (e.g., they were required to remember 4 items and rated 4). Divergent responses were scores where the rating diverged from the number of items the participant was required to remember (e.g., required to remember 4 items, reported remembering 2 items, divergence score for trial = 2). This showed a strong negative correlation between confidence ratings and divergence score ($r = -.737$, $p < .001$), suggesting the lower the divergence between the number of to-be-remembered items and the number of items in mind, the greater the confidence participants had in their VWM performance.

#### 4.3.2.2. Proportion correct between low vs. high vividness trials and non-divergent vs. divergent quantity trials

To investigate whether individual's subjective ratings reflected VWM accuracy, two paired sample t tests were conducted to examine the difference in proportion correct between trials rated with high vividness and low vividness and non-divergent and divergent quantity ratings, respectively. High vividness ratings were trials where the participant rated either 3 (moderate image) or 4 (strong image/almost like perception) and low vividness ratings were trials where the participant rated either 1 (almost no image) or 2 (weak image). Non-divergent quantity ratings and divergent quantity ratings were as described above. Firstly, there was no significant difference between proportion correct in high vividness trials (*M* = .73, *SD* = .15) and low vividness trials (*M* = .67, *SD* = .12) (t(16) = 1.51, *p* = .152, *d* = .37). For the quantity ratings analysis, one participant was excluded because none of their trials were divergent, and another participant was excluded as none of their trials were non-divergent. There was a significant difference between proportion correct in non-divergent ratings (*M* = .77, *SD* = .07) and divergent ratings (*M* = .68, *SD* = .19) (t(14) = 2.21, *p* = .04, *d* = .57), which showed greater accuracy in non-divergent trials compared to divergent trials.

4.3.2.3. Relationship between proportion correct and subjective MI ratings as a function of set size

Given the set size effect observed in both proportion correct and ratings reported in **Section 4.3.1.1.** and **Section 4.3.1.2.**, respectively, further analyses into the relationship between proportion correct and ratings were conducted. First, Spearman's correlations were conducted to examine the relationship between proportion and rating at each set size. As the analyses above indicate only an effect of set size in ratings and proportion correct; precision, instruction and attended side were collapsed across to retain power in the following analyses. Vividness ratings were significantly and positively associated with proportion correct in 1 item trials ($r_s$ = .578, *p* = .015), however vividness ratings and proportion correct were not significantly associated in 2 item trials ($r_s$ = .143, *p* = .585) or 4 item trials ($r_s$ = .010, *p* = .974).

For the quantity ratings analysis, the divergence score was included. Quantity divergence was not associated with proportion correct in 1 item trials ($r_s$ = -.427, *p* = .088), 2 item trials ($r_s$ = -.369, *p* = .144) or 4 item trials ($r_s$ = -.327, *p* = .200). Taken together, the findings suggest participants have relatively poor insight into the visual quality (vividness rating) of representations held in VWM and the number of visual items (quantity rating) in representations held in VWM, except for visual quality (vividness) at the smallest set size (1 item).

4.3.2.4. CDA in high vs. low vividness ratings and non-divergent vs. divergent quantity ratings

To examine CDA between rating type at each set size and instruction, an ANOVA was planned with grand-averaged CDA as the dependent variable and rating (high vividness, low vividness, non-divergent quantity, divergent quantity), set size (1 item, 2 items, 4 items), instruction (quality-focused, capacity-focused) and attended side (left, right) as the within-subject factors. However, as the conditions were based on participant responses, there was at least one condition per participant where there were no responses (e.g., some participants did not rate any 4 item trials as high vividness). The descriptive statistics for the number of trials per type of rating (high vividness, low vividness, non-divergent, divergent ratings) per instruction (capacity-focused, quality-focused) and attended side (left, right) condition are reported in Tables A.3.5-A.3.7 in the **Appendix**. Therefore, an ANOVA was conducted for vividness ratings and quantity ratings collapsed across all conditions except vividness (number of high vividness responses: $M$ = 91, $SD$ = 43, *range* = 34 to 151; number of low vividness responses: $M$ = 68, $SD$ = 46, *range* = 4 to 140) and quantity (number of non-divergent responses: $M$ = 115, $SD$ = 35, *range* = 65 to 174; number of divergent responses: $M$ = 50, $SD$ = 35, *range* = 0 to 110) respectively. The vividness rating ANOVA included a within-subject factor of vividness (high, low), which revealed no main effect of vividness ($F(1,16)$ = 1.38, $p$ = .258, $\eta_p^2$ = .08). The quantity ratings ANOVA included within-subject factors of divergence (non-divergent, divergent). Similarly, to the vividness ANOVA, there was no main effect of divergence ($F < 1$).

4.3.2.5. Relationship between CDA and ratings as a function of set size

To assess the relationship between CDA and subjective MI ratings, separate correlations were conducted for vividness ratings and quantity ratings. For vividness ratings, the CDA dependent variable was computed as the difference between grand-averaged CDA for 1-item trials and 2-items trials per participant, given that vividness is expected to be more prominent in smaller set sizes. The vividness ratings dependent variable consisted of the mean vividness ratings for 2-item trials per participant. This is based on the logic that if vividness ratings map onto the number of items in mind as indexed by CDA, there should be a positive association between vividness ratings in 2-item trials and the difference in CDA between 1- and 2-item trials, i.e., the greater the set size effect in CDA, the higher the vividness rating. However, there was no relationship between the difference between CDA in 1-item and 2-item trials and vividness ratings in 2-item trials ($r_s$ = -.314, $p$ = .220).

For quantity ratings, the CDA dependent variable was computed as the difference between grand-averaged CDA for 1-item trials and 4-item trials. The quantity ratings dependent variable consisted of the mean quantity rating for 4-item trials. As above, this is based on the logic that if quantity ratings map onto the number of items held in mind as indexed by CDA, there should be a positive association between quantity ratings in 4-item trials and the difference between CDA between 1-item and 4-item trials, i.e., the greater the set size effect in CDA, the more items the participant reports holding in mind. However, there was no relationship between the difference between CDA in 1-item and 4-item trials and quantity ratings in 4-item trials ($r_s$ = .302, $p$ = .239).

#### 4.3.2.6. ADAN in high vs. low vividness ratings and non-divergent vs. divergent quantity ratings

To examine how ADAN differed in high vs. low vividness trials and non-divergent vs. divergent quantity trials, two ANOVAs were conducted on vividness ratings and quantity ratings with grand-averaged ADAN as the dependent variable and either vividness (high vividness, low vividness) or divergence (non-divergent, divergent) as the within-subject factors, respectively. One participant was excluded from the vividness ratings ANOVA as the participant didn’t rate any trials as low vividness. Three participants were excluded from the quantity ratings ANOVA as there were no divergent ratings for those participants (descriptive statistics for number of trials are reported in Table A.3.4 in the **Appendix**). In the vividness ANOVA, there was no main effect of vividness ($F(1,15) = 1.09$, $p$ = .314, $\eta_p^2$ = .01). Similar results were found in the quantity ratings ANOVA; there was no main effect of divergence ($F(1,13) = 2.69$, $p$ = .124, $\eta_p^2$ = .17). Grand-averaged waveforms for ADAN for high vs. low vividness ratings and non-divergent vs. divergent quantity scores are depicted in Figure 4.12.

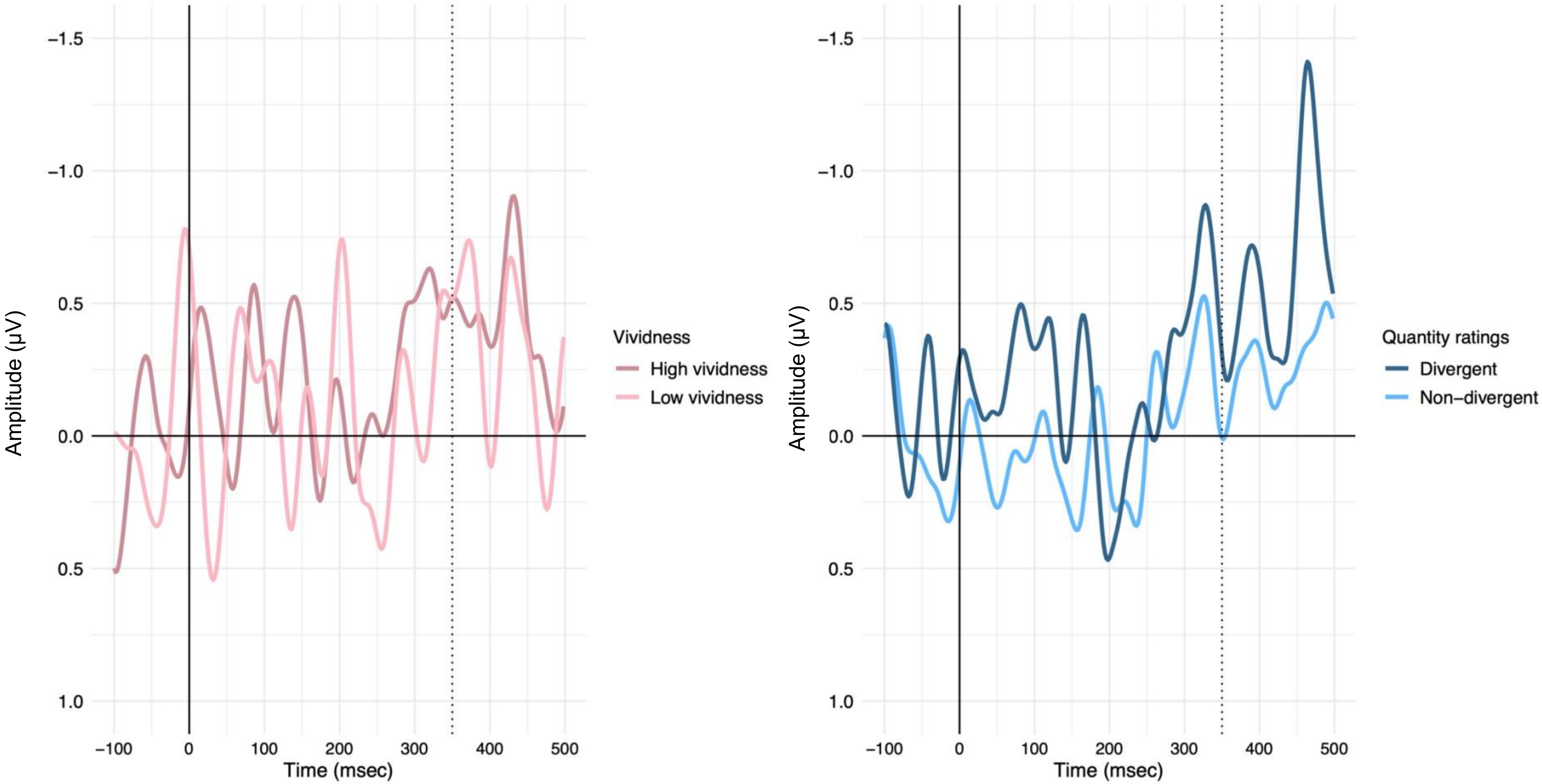



*Figure 4.12:* Grand-averaged waveforms of ADAN for high and low vividness (left) and divergent and non-divergent (right) ratings. Cue onset is at 0-200ms and dotted line at 350msec added for reference (ADAN amplitude calculated as mean amplitude between 350ms and 500ms)

## 4.4. Discussion

Previous literature, and the **Chapters 2** and **3** of this thesis, have made inferences regarding how MI is recruited in VWM by associating performance on separate MI and VWM tasks. To directly examine how MI, and explicitly how visual representations, are recruited in VWM, the study presented in this chapter examined MI *within* a VWM task. The first aim was to characterise the visual precision and capacity of VWM maintenance with respect to CDA, proportion correct and subjective ratings of the VWM maintenance period. The second aim was to establish the metacognitive link between the subjective sensory experience of MI and VWM. With reference to the first aim, previous evidence for greater accuracy, as measured by proportion correct, at smaller set sizes compared to larger set sizes was supported in this study. In line with predictions, the precision effect found in previous research (greater proportion correct in coarse vs. fine trials; Machizawa et al., 2012) was modulated depending on whether participants were following a capacity-focused or quality-focused instruction: greater proportion correct was observed between coarse and fine trials in the quality-focused instruction condition only (although, there was only a trend towards significance). Proportion correct also differed between right attended and left attended trials. Firstly, in coarse precision and quantity rating conditions only, proportion correct was greater in right attended compared to left attended trials. Second, it was found that in the capacity-focused, quantity rating block, proportion correct was greater proportion correct was indicated in the right attended trials compared to left attended trials in the 4-item condition only. Again, these two interactions were supported only by a trend towards significance and are therefore interpreted cautiously in the main discussion. With respect to subjective ratings, vividness ratings were higher in smaller set sizes (1-item trials) compared to larger set sizes (4-item trials) and quantity ratings increased with increasing set size as predicted. However, it was expected that instruction would modulate vividness ratings in that vividness ratings would be higher in quality-focused blocks compared to capacity-focused, but there was no difference between instruction blocks. Moreover, quantity ratings were expected to be higher when participants were following a capacity-focused instruction, however there was no effect of instruction.

With regards to neural correlates of VWM maintenance, the established finding that CDA indexes number of items held in mind during VWM up to 4 items was replicated in that greater CDA amplitude was observed in larger set sizes compared to smaller set sizes. There was a trend for the modulation of CDA by set size and metacognitive rating. This revealed a set size effect for both vividness and quantity ratings but the pattern of results was different; in the vividness ratings, there was a step change between 2 and 4 items, but the step change in quantity ratings blocks was between 1 and 4 items. However, it is important to note that this was only a trend towards significance and there were no significant differences between the vividness and quantity ratings at each set size in the post hoc comparisons. Attended side also modulated CDA amplitude and the pattern of results was different to the pattern reported in proportion correct. In the capacity-focused blocks, greater CDA amplitude was observed in coarse compared to fine trials in right attend trials but not left attend trials; this was unexpected given that precision (fine and coarse trials) was not cued in this experiment.

Next, to address the second part of the first aim, individual differences in the relationship between subjective MI ratings and behavioural and neural correlates of VWM were examined. Correlations between vividness ratings and quantity ratings, respectively, suggested as vividness ratings increased, confidence in proportion correct increased, and as quantity divergence decreased, confidence in proportion correct increased. There was no difference in proportion correct between high vividness and low vividness trials, however significantly higher proportion correct was observed in non-divergent quantity ratings compared to divergent ratings. This suggests that while individuals have poor insight into the vividness of MI during VWM, they have better insight into the number of items retained during VWM. There was also no difference in CDA amplitude between high and low vividness trials, and non-divergent and divergent quantity rating trials, respectively. Further investigation looked at associations between proportion correct and CDA and subjective MI ratings. This revealed a positive association between proportion correct and vividness ratings in 1-item trials only, and there was no relationship between divergence score and quantity ratings for each set size. Moreover, there was no significant association between the difference in CDA between 2- item and 1-item trials and vividness ratings, and no significant association between the differences in

CDA between 4-item and 1-item trials and quantity ratings. Finally, there was no difference in ADAN amplitude between high and low vividness ratings and non-divergent and divergent quantity ratings. Overall, the findings suggest a disconnect between behavioural and neural correlates of VWM. The findings are discussed in turn below.

### 4.4.1. Behavioural and neural correlates of VWM maintenance are modulated by both instruction and subjective rating type

Previous findings were replicated in that proportion correct was greater in smaller set sizes compared to larger set sizes (Machizawa et al., 2012; Machizawa et al., 2020). Findings were also extended by demonstrating evidence for modulations of proportion correct dependent on instruction and precision such that proportion correct for coarse trials was greater that fine trials in quality-focused blocks only, as expected. Caution in interpreting this result should be applied given that there was only a trend towards significance. Therefore, while findings are interpreted in the context of previous literature, replication is required to either confirm or deny these inferences. Namely, this suggests that participants were prepared to retain visual information with high precision, even in the easier (coarse precision) trials, when following a quality-focused instruction. Previous studies have shown greater proportion correct in coarse compared to fine trials and quality-focused fine trials compared to capacity-focused coarse trials, respectively (Machizawa et al., 2012; Machizawa et al., 2020). However, in the paradigm presented in this chapter, participants were only aware of the instruction, either quality-focused or capacity-focused, and fine and coarse trials randomised within the block. Therefore, this finding further supports the suggestion that individuals exert wilful control over the precision of visual representations, as instructed, and this in turn influences their performance (Machizawa et al., 2012; Machizawa et al., 2020; Zhang & Luck, 2008). With respect to subjective ratings, vividness ratings were higher at smaller set sizes and quantity ratings increased with increasing set size. Moreover, there were no effects of instruction on vividness or quantity ratings. This suggests that the type of instruction did not modulate individuals' subjective experience of the number of items held in mind, which is perhaps not surprising given that the ratings are subjective in nature.

The effects of attended side in proportion correct and CDA amplitude are notable. Proportion correct was significantly greater in right attend trials compared to left attended in the capacity-focused condition only, alongside a trend towards significance for an interaction which revealed in 4-item trials proportion correct was greater in right attend compared to left attend trials in the capacity-focused, quantity rating block only. Comparatively, significantly greater CDA amplitudes were indexed in coarse compared to fine trials in the right attend but not left attend trials in the capacity-focused blocks. It is important to note that hemispheric differences were not hypothesised. Moreover, attended side was included in the ANOVAs as a quality control measure given that CDA is calculated as the difference between contralateral and ipsilateral electrodes to the attended stimuli. However, previous research has examined hemispheric differences in VWM as indicated by CDA, therefore the results are discussed in the context of this research. Namely, Machizawa et al. (2020) report that behavioural performance and CDA amplitudes in their quality-focused instruction condition (fine trials only) were associated with the grey matter volume in the right parietal cortex whereas behavioural performance and CDA amplitudes in their capacity-focused condition (coarse trials only) were associated with grey matter volume in the left lateral occipital cortex. The findings presented here, that are specific to the largest set size when participants were required to rate the number of items in mind (quantity rating) and were following a capacity-focused instruction, support the indication of left hemispheric specialisation of VWM capacity. Previous findings are therefore extended by demonstrating that this effect was only present when participants were required to rate quantity, suggesting that the participants expectation to rate the number of items in their representation impacted performance.

The CDA findings are more difficult to disentangle. The finding of significantly greater amplitude in coarse trials compared to fine in the right attend trials only is perhaps not entirely surprising as it is partially in line with an association between coarse (capacity-focused) performance and left lateral occipital volume in Machizawa et al.’s (2020) study, although in their study coarse precision was cued. Therefore, the finding that there is a difference between coarse and fine trials is unexpected in that participants were not cued for the precision (fine, coarse) modulation in the current study. Given that the 3- and 4-way interactions include individual conditions with limited number of trials per condition, it is not possible to make general conclusions regarding hemispheric differences in CDA based on these findings and further research is warranted.

Findings regarding modulation of CDA dependent on ratings conditions is in line with an executive-attention account of VWM. Conceptualised by Braver et al. (2007), proactive control describes the finding that individuals prepare to a response in advance of a stimulus based on known information, which is indexed by modulation of neural correlates in VWM. This is supported in Machizawa et al.’s (2012) findings demonstrating proportion correct and CDA were modulated dependent on whether participants expected fine or coarse precision trials. Previous evidence was extended in this chapter by demonstrating that the CDA set size effect was modulated depending on whether participants were instructed to rate the vividness of their representation compared to the quantity of items in their visual representation. Prior to exploring the interpretations of this finding with respect to previous literature, it is important to highlight that there was only a trend towards significance for this interaction. When participants were prepared to rate the vividness of items in mind, the set size effect was present but not between 1 item and 2 item conditions and when participants were prepared to rate the quantity of items in mind the set size effect was present but not between the 2 item and 4 item conditions. This suggests that the participants expectation to rate either quantity or vividness modulated their memory consumption such that the asymptote was reached at a smaller set size in quantity rating trials (2 items) as opposed to the larger set size in vividness rating trials (4 items). This is therefore in line with a proactive control account of VWM (Braver et al., 2007) and extends previous findings demonstrating expectation of precision can modulate CDA (Machizawa et al., 2012) to demonstrate that individuals exert control over the content of VWM dependent on the type of metacognitive rating they expect to report in that block. In addition, while there was no difference between CDA amplitude in 1-item vividness rating trials compared to 1-item quantity rating trials, it is important to acknowledge that CDA amplitudes in the 1-item vividness trials appear to be greater than previously reported 1-item trials. For instance, the mean CDA amplitude reported in Vogel and Machizawa (2004) is around -.70 whereas the mean CDA amplitude reported in this study in the vividness rating trials is around -1.10 (as depicted in Figure 4.9). This further implies that the requirement to rate the vividness of visual representations modulated the wilful control exerted to retain visual information in mind, however, due to the trend towards significance, further research powered to detect 3- and 4-way interactions is required to support this conclusion.

The findings therefore extend current understanding to demonstrate that individuals not only respond to cue-related expectations within trials (as demonstrated in Machizawa et al., 2012), but also the blocked instructions and the requirement to rate information in mind, and these nuances in the control over the visual representations are indexed by CDA. It would be informative to assess how this differs dependent on variability in VWM capacity. Research has shown that individuals with high VWM capacity exert greater proactive control, with regard to applying cue information to prepare responses, compared to individuals with low VWM capacity (Redick, 2014). However, given the sample size restrictions in this study, examining between-group differences of high and low VWM capacity is beyond the remit of this chapter.

### 4.4.2. Limited metacognitive link between MI and VWM

While the instruction and metacognitive ratings conditions modulated proportion correct and CDA respectively, metacognitive ratings do not appear to reflect the precision at which visual information is held in mind, contrary to hypotheses. Although proportion correct was higher in the high vividness rating trials compared to low vividness rating trials, this difference was not significant. Vividness was significantly correlated with proportion correct in 1-item trials, but not in 2- or 4-item trials. Moreover, there were no differences between CDA amplitude between high vividness rating trials and low vividness rating trials and no significant association between the CDA set size effect and vividness ratings. It should first be noted that although there are no significant differences, the means indicated greater CDA in low vividness compared to high vividness and in divergent compared to non-divergent; the opposite to what was expected. This might suggest a compensatory mechanism; if individuals *expect* that they will only maintain a visual representation of low vividness and if individuals expect to not be able to hold the correct number of items in mind, this might result in greater memory consumption/proactive control as indexed by CDA. Clearly, further investigation is needed to test this hypothesis.

The findings are important because they suggest that although individuals can flexibly exert control of the visual precision and capacity of visual information held in mind, they appear to have relatively poor insight into both the visual quality and capacity of visual representations held in mind during VWM. In Pearson & Keogh's (2019) review, they argued that individual differences in the neural correlates of VWM may be dependent on the types of strategies recruited in VWM, i.e., imagery strategies vs. propositional-symbolic strategies, and that measuring strategies recruited in VWM tasks might explain these individual differences. The study presented in this chapter directly addresses this proposition by measuring trial-by-trial subjective ratings of MI within a VWM task. However, the evidence presented here demonstrates a distinction between individual differences in self-reported, subjective ratings/MI strategies and individual differences in the precision and capacity at which visual information is held in mind. Firstly, propositional/verbal strategies are unlikely in this task given the very short stimuli presentations (200ms) and delay period (1400ms); therefore, this confound can be ruled out. Moreover, the modulations in proportion correct and CDA amplitude depending on instruction and type of rating demonstrate that individuals have flexible control over the precision and capacity at which visual information is held in mind, as discussed in detail above. Instead, the evidence presented here suggests a dissociation between the subjective sensory experience of MI and how visual information is maintained VWM. There are just two other known studies that have examined the relationship between behavioural outcomes in VWM and MI (Keogh & Pearson, 2011; 2014). Findings show that MI sensory strength was positively associated with VWM capacity at set size 3 (Keogh & Pearson, 2014) and only VWM performance in those with high MI sensory strength was disrupted by background luminance manipulations (Keogh & Pearson, 2011; 2014). While in those studies it was argued that individual with stronger MI recruit MI strategies in VWM, the findings presented in this chapter call into question whether assessing subjective strategies in VWM is akin to behavioural and neural indices of the visual precision and number of visual items maintained in VWM.

The relationship between quantity ratings and the behavioural and neural correlates of VWM maintenance are more complex. Firstly, the increases in CDA amplitude were more prominent between the largest set sizes and the smallest set size in the quantity ratings blocks and more prominent between the largest and two smallest set sizes in the vividness ratings blocks. However, keep in mind that this was indicated in a non-significant interaction with only a trend towards significance. Second, significantly greater proportion correct was observed in non-divergent trials compared to divergent trials. Therefore, this suggests that participant's *expectation* to rate the number of items in mind modulated their memory consumption/proactive control, as indexed by CDA amplitude modulations. However, there was no difference in CDA between non-divergent compared to divergent trials and no relationship between the CDA set size effect between 1-item and 4-item trials and quantity ratings for 4-item trials. This therefore suggests that individuals may have reasonably better insight into the quantity of items in mind compared to vividness of representations. Furthermore, based on the evidence reviewed in **Chapter 1, Section 1.2.1.2.**, it was hypothesised that ADAN would depict the functional role of frontal regions in MI in that ADAN would be modulated by vividness ratings and quantity ratings, respectively. However, this was not supported in that there was no difference between ADAN amplitudes in high and low vividness ratings or between non-divergent and divergent ratings, respectively. The grand-averaged waveform for ADAN amplitudes suggests greater ADAN amplitude in divergent trials compared to non-divergent trials, however this difference was not significant. It could be speculated that this reflects a difference in approach to the task, i.e., those who think they are able to remember less items are more attentive prior to the sample array in an attempt to remember more items. However, further investigation is required to test whether this conclusion is supported. Moreover, the findings regarding the metacognitive link between MI and VWM are somewhat limited due infrequent responses. For example, some participants didn't rate any 4-item trials as low vividness, therefore it was not possible to test the relationship between individual differences in ratings and CDA for each set size or instruction, for instance. Future studies sampling participants based on low vividness, high vividness, non-divergent and divergent ratings would be useful to further examine individual differences. Overall, the link between the metacognitive sensory experience of MI and behavioural and neural correlates of VWM is tentative at best.

### 4.4.3. General considerations

It is important to consider potential methodological constraints. Previous evidence has suggested a relationship between saccades and MI in that participants tend to make similar gaze patterns when imagining a previously viewed stimulus as they do when viewing a stimulus, known as the “looking at nothing” effect (Brandt & Stark, 1997; Johansson & Johansson, 2014). Given the nature of EEG data, trials with saccades present artefacts which must be removed prior to analysis. While ICA was conducted to retain as many trials as possible and the number of trials retained per participant was > 75% in this chapter, this is an important consideration given that high MI trials may have been rejected due to saccade artefacts. However, a recent study examining gaze patterns during MI found that gaze patterns during MI trials were not associated with vividness of MI as measured by the VVIQ (Gurtner et al., 2021). Therefore, it appears that it is unlikely that rejection of saccade trials would have influenced results examining the metacognitive link between MI and VWM in this study. Future research examining gaze patterns alongside subjective ratings of MI within a VWM task would further elucidate this relationship.

It is also notable that participants appear to rarely rate at either end of the rating scales. For example, individuals rarely report having 4 items in mind in the quantity ratings. The fact that the vividness rating scale ranged from 1-4 and the quantity rating scale ranged from 0-4 could have been confusing. However, the vividness rating scale was chosen as so in line with previous studies (Pearson et al., 2011; Dijkstra et al., 2017b). This is the first study to adopt a quantity rating scale and it appears individuals are reluctant to rate at either end of the scale. One previous study has used a continuous scale for rating vividness using a sliding bar (Dijkstra et al., 2020), however responses broadly fell into the 1-4 category ratings and were therefore binned as such. Further research is required to test whether ratings are distorted by the Likert-scale.

In addition, it is important to recognise the limited sample size. Twenty-three participants were recruited, which is in line with previous studies demonstrating robust CDA effects in precision and capacity in VWM (Machizawa et al., 2012; Machizawa et al., 2020). However, due to exclusion, only 17 participants remained in the final sample. Moreover, the instruction condition was added followed piloting with the aim of reducing the influence of difficulty and replacing the expected condition of precision (fine, coarse). This rendered 5-factor ANOVAs with low power. Small sample sizes are a common issue in neuroimaging studies given the resource and time constraints associated with this research. Recently, it was suggested that only 30-50 trials are needed to detect the presence of CDA but for differences between set size up to 400 trials per condition could be required (Ngiam et al., 2021). While this is informative, up to 400 trials per condition is practically very difficult as this would lead to lengthy experiments and therefore participant fatigue and boredom, which would distort the data. Clearly, it is important to strike a balance in methodological design and to take sample size and trial numbers into account when drawing conclusions on analyses of both CDA and ADAN.

### 4.4.4. Conclusions

The evidence presented in this chapter provides important methodological and theoretical contributions to current understanding of the metacognitive link between MI and VWM. Previous findings with respect to precision and capacity of VWM maintenance, as indexed by proportion correct and CDA amplitude, were firstly extended as it was demonstrated that both were modulated by instruction, type of subjective MI ratings and attended side. Proportion correct was found to be greater in right attend to left attend trials in the 4-item condition of the capacity-focused, quantity rating block only. An effect of attended side was also demonstrated in that greater CDA amplitudes were evidenced in coarse compared to fine trials in the right attend trials in the capacity-focused blocks only. These findings are in line with previous evidence for left hemispheric specialisation of VWM capacity. Moreover, the CDA set size effect showed a different pattern of results in the vividness ratings blocks compared to the quantity ratings blocks in that there were differences between all conditions apart from between 1-item and 2-item trials in the former, and differences between all conditions apart from between 2-item and 4-item trials in the latter (however, there was only a trend towards significance for the interaction between set size and type of rating). These novel findings are in line with a proactive control account of VWM and support the notion that individuals can exert wilful control over the precision and capacity at which information is held in mind.

Importantly, the prediction of a metacognitive link between subjective MI ratings and VWM was not supported. Firstly, there were no differences between high and low vividness trials in proportion correct, CDA amplitude or ADAN amplitude. While proportion correct was higher in non-divergent quantity ratings compared to divergent quantity ratings, there was no difference in CDA amplitude or ADAN amplitude. Therefore, contrary to hypotheses, individuals appear to have poor insight into the visual precision and capacity of representations held in VWM. Rather than providing a novel method for quantifying the role of MI in VWM using metacognitive ratings, this study importantly highlights the disconnect between subjective ratings and the visual precision and quantity of VWM. This has important methodological implications for examining how individual differences in MI support VWM and provides novel contributions demonstrating a dissociation between the subjective sensory experience of MI and behavioural and neural correlates of VWM.

# Chapter 5: General discussion

## 5.1. Thesis overview

The experimental studies presented in this thesis provide novel theoretical and practical contributions to the understanding of how individual differences in MI, VWM and attention interact in adults, typically developing (TD) children and children with ADHD. To address the gaps in the literature, **Chapter 2** firstly provides a detailed account of how visual mental images are generated (image generation), maintained (image maintenance) and manipulated (mental rotation and image scanning) in primary school years (children aged 6-11 years) and adulthood using a novel battery of tasks designed to tap into the visual precision of mental images. Second, the associations between the components of MI are investigated to establish how components of MI are related. Third, the relationship between components of MI, VWM maintenance and manipulation and attention control is investigated in children and adults. To establish a full picture of how MI presents alongside VWM in both typical development and in children with ADHD, **Chapter 3** adopts a case-control design and individual differences approach to characterise abilities in the context of a component model of MI and VWM. This chapter provides a much-needed account of how MI presents alongside VWM in children with ADHD compared to TD children, as well as highlighting important individual differences in these abilities. Finally, while current research has examined how measures of MI relate to measures of VWM, research is yet to investigate the role of MI within a VWM task paradigm even though the argument that individual differences in the recruitment of MI might underpin individual differences in VWM performance has been made (Pearson & Keogh, 2019). To address this gap in the literature, a novel VWM paradigm was introduced in **Chapter 4** to examine how individual differences in visual quantity of items in mind and visual quality of MI impacted behavioural and neural correlates of VWM capacity and precision. Given the recent acknowledgement of a functional role of both visual and frontal regions in MI (Spagna et al., 2020), individual differences in ADAN dependent on the visual quality and number of items in MI were also assessed for the first time in this chapter.

The general discussion chapter provides a summary of the main results reported in the experimental chapters and discusses the theoretical and practical implications of these findings. Limitations and considerations for future research are then outlined, followed by the concluding remarks of the thesis.

## 5.2. Summary of results

A principal aim of **Chapter 2** was to adopt a component model of MI to examine the extent to which a depictive theory of MI is recruited in childhood. The format of visual mental images has been investigated extensively in adults primarily using neuroimaging methods (reviewed in detail in **Chapter 1, Section 1.1.1.**), which has demonstrated that individuals can generate visually depictive mental images in the absence of sensory input (e.g., Kosslyn & Thompson, 2003; Lee et al., 2012, Naselaris et al., 2015). The use of the term depictive images refers to the mapping of spatial coordinates of a previously viewed stimulus to retinotopically organised areas of the visual cortex (Ganis, 2013; Kosslyn et al., 2006). However, the depictive theory of MI more generally refers to the ability to generate and maintain visual images. Therefore, by modulating the visual precision at which stimuli is generated and held in mind in behavioural studies (Kosslyn et al., 1990), it is possible to detect evidence for a depictive theory of MI. Little is currently known regarding the extent to which a depictive theory of MI is supported in childhood.

Only two studies to date have examined how image generation and image maintenance develop throughout childhood and vital questions regarding support for a depictive theory remain. In the first study, tasks rendered too difficult led the youngest age group tested (5 years of age) to be excluded from analysis (Kosslyn et al., 1990). The next age group tested in this study was 8 years old, therefore understanding of how MI develops in early primary school years was missing. A more recent study introduced more age-appropriate paradigms; however, the generation and maintenance tasks involve remembering visuo-spatial locations of known objects (Wimmer et al., 2015), which likely taps into visuo-spatial imagery rather than the visual precision of mental images. While participants could be referring to mental images in these tasks, the recruitment of verbal strategies cannot be ruled out. Therefore, novel tasks were designed to examine the extent to which children generate (image generation) and maintain mental images (image maintenance) of high visual precision. Firstly, in **Chapter 2**, children of all ages demonstrated the ability to generate and maintain mental images of high precision, as evidenced by faster RTs, compared to low precision responses. With regard to development, it is notable that different patterns of results were observed in image generation and image maintenance, and these also differed in comparison to previous literature (Wimmer et al., 2015). In their image generation task, Wimmer et al. (2015) found adult-like abilities were observed in children age 8 years when participants were required to detect the location of the known object (first experiment) and in children age 6 years when participants were required to drag and drop the object to the correct location (second experiment). The second experiment was argued to be more sensitive to the precision of mental images. In their image maintenance task, Wimmer et al. (2015) found adult-like abilities in children age 8 years in their first experiment and their second experiment. The findings in **Chapter 2** suggest the ability to generate and maintain mental images of high precision develop later than as suggested in Wimmer et al. (2015); adult-like abilities were reached at around 8-9-years in image generation and abilities were still developing into later primary school years (10-11-years) in image maintenance. In addition, Wimmer et al.'s (2015) findings showed greater accuracy and faster RT in image maintenance compared to image generation in all age groups, however the study presented in **Chapter 2** found this pattern of results was only present in adults. Thus, it was interpreted that while the ability to generate and maintain visuo-spatial mental images may develop earlier (as evidence in Wimmer et al., 2015), the ability to generate and maintain visual images of high precision reaches adult-like levels in later childhood. Given that the ability to generate and maintain mental images of high precision from age 6 was found in this study, this suggests support for a depictive theory of MI in children as young as 6 years old.

As part of the component model of MI, mental rotation and image scanning were also assessed. Previous evidence for mental rotation ability and image scanning ability was replicated in that children from age 6 years demonstrated the linear increasing time-degree of rotation effect (Estes, 1998; Frick et al., 2013; Marmor, 1975; Wimmer et al., 2017) and linear increasing time-distance effect (Wimmer et al., 2016; Borst & Kosslyn, 2011), respectively. Current understanding regarding the precision at which visual mental images are recruited in the image scanning task was extended by demonstrating that distance was underestimated and that this deviation from actual distance increased with increasing distance. This pattern of results was present in children of all ages and in adults and the same pattern of results was found in perception trials. These findings are in line with findings of distance estimation in visual perception (Norman et al., 2016) and thus further support evidence for shared mechanisms between MI and visual perception (Dijkstra, Bosch, et al., 2019) and extend this to childhood. These findings are the first to highlight both children and adults make similar errors in estimated distance in the absence of sensory input compared to visual distance perception, further supporting evidence for a depictive theory of MI from age 6 years.

The second aim of this chapter was to examine how components of MI were associated throughout development. Findings showed that none of the components of MI were associated with one another in childhood. Importantly, this provides evidence in support of a separable-component model of MI in childhood (Kosslyn et al., 1990). It is firstly notable that there was a significant association between image generation and image maintenance RT in the child data correlation matrix, but this positive association was not significant between the image generation and image maintenance accuracy measures. While it might be suspected that the RT association is dependent on development in processing speed, a replication study is required to warrant this conclusion. In adulthood, **Chapter 2** demonstrated evidence for a relationship between the ability to maintain mental images (image maintenance) and the ability to manipulate mental images (mental rotation). This is in line with previous evidence for the suggestion that visual representations are recruited in spatial transformations such as in mental rotation (Hyun & Luck, 2007; Prime & Jolicoeur, 2010). Therefore, it is altogether not surprising that the two components are integrated in adulthood, and it is notable that they appear to develop separately throughout the primary school years.

The third and final aim of **Chapter 2** was to establish how components of MI relate to VWM maintenance and manipulation and attention control. While previous studies have speculated on the role of VWM in MI abilities in childhood (Wimmer et al., 2017), this has not been tested directly in children until now. Correlational analyses firstly revealed that there were no significant associations between each of the components of MI and either VWM maintenance or VWM manipulation in primary school aged children or adults. This is somewhat surprising given the overlap in the theoretical definitions of MI and VWM (e.g., Cornoldi et al., 2003) and previous evidence for shared visual representations between MI and VWM (Albers et al., 2013). Instead, this evidence points towards individual differences in the types of strategies recruited in MI and VWM tasks. Thus, while it might be assumed that both involve the recruitment of visual representations, this might not be the way that all individuals approach this task. Findings suggesting that only "good imagers" appear to recruit MI in VWM supports this suggestion (Keogh & Pearson, 2011; 2014). Thus, while a relationship between MI and VWM was initially expected at the group level, it might be that there are distinctions between how children and adults with varying levels of MI ability approach VWM tasks.

Interestingly, the relationship between attention control and components of MI varied between children and adults. In children, there were no associations between components of MI and attention control. However, in the adult group, there was evidence for a relationship between attention control and image generation, image maintenance and mental rotation. These findings are tentatively reported as the *p* values were just above the highly conservative, Bonferroni multiple comparison cut-off. Previous investigation into the relationship between MI and attention is limited. On the other hand, the internal focus of attention is central to theoretical frameworks of VWM with respect to maintaining stable visual representations, which is argued to develop with improved attention control throughout childhood (Cowan, 2016; Shimi et al., 2014). In **Chapter 2**, MI abilities and attention control appear to develop separably and there is an indication that adults recruit attention control in MI.

In the light of evidence for heterogeneity of working memory abilities in children with and without ADHD (e.g., Campez et al., 2020) alongside evidence for a positive relationship between MI and VWM in adults (Keogh & Pearson, 2011; 2014), it is surprising that MI abilities have not been examined in children with ADHD. This thesis presents the first investigation characterising abilities in the components of MI alongside maintenance and manipulation measures of VWM using both a case-control design and an individual differences approach in both TD children and children with ADHD. First, between-group analyses revealed children with ADHD demonstrated a typical pattern of performance in each component of MI, except for shallower slopes in RT in the image scanning task, compared to TD children of the same age. Specifically, children with ADHD showed the same pattern as TD children and the broader sample of children in **Chapter 2** in the image generation and maintenance tasks suggesting that this group can generate and maintain mental images of high visual precision. It is also notable that the ADHD group in this chapter presented with typical levels of ability in mental rotation accuracy, contrary to previous indications of poorer performance compared to TD controls (e.g., Silk et al., 2005; Vance et al., 2007). Children with ADHD showed the linear time-degree of rotation effect, which is evidenced in TD children in this chapter, the broader sample of primary school children in **Chapter 2**, and previous literature (Estes, 1998; Frick et al., 2014; Möhring et al., 2016; Wimmer et al., 2017). Further investigation of mentally representing varying distances in image scanning revealed children with ADHD showed the same pattern of results as TD children; deviation from actual distance increased with increasing distance. Overall, children with ADHD demonstrated broadly typical patterns of abilities and age-appropriate levels of ability in each component of MI and VWM.

Profiles of ability were examined in MI and VWM in between-group analyses and data-driven individual differences analyses. This revealed at a group level that the profile of MI and VWM abilities in children with ADHD did not differ from TD children of the same age. Additionally, a latent profile analysis found evidence for five distinct profiles of ability that were transdiagnostic in nature; individual differences in MI and VWM were not syndrome-specific to ADHD. This is a vital finding considering the current understanding of the neuropsychological profile of ADHD. For several decades, research endeavoured to determine neuropsychological deficits in ADHD (Barkley, 1997; Castellanos et al., 2006; Nigg et al., 2005). This research recruiting case-control designs led to the suggestion that children with ADHD present with VWM impairments (Kasper et al., 2012). However, more recent research using individual differences approaches to analysis has revealed extensive variation in VWM abilities in that both children with ADHD and TD children present with both poor, moderate, and high VWM abilities (Campez et al., 2020). The findings from the latent profile analysis partially support previous research in that children with ADHD fell into the moderate (profile 2, 4 and 5) and high (profile 1) VWM ability profiles. No children with ADHD fell into the profile characterised by relatively low VWM maintenance, however given the limited sample size, this likely suggests this particular ADHD group had relatively high VWM abilities compared to what might be expected based on previous research (e.g., Kasper et al., 2012). Beyond the evidence for transdiagnostic profiles of abilities revealed in the latent profile analysis, this investigation has provided further evidence for dissociated MI and VWM abilities. Specifically, some children presented with relatively high mental rotation abilities alongside high VWM abilities (profile 1) and moderate VWM abilities (profile 4), respectively, but others presented with low mental rotation abilities alongside moderate VWM abilities (profile 2) and moderate mental rotation abilities alongside relatively low VWM abilities (profile 3). In short, this suggests that mental rotation abilities are dissociated from VWM abilities in children and implies there are individual differences in how visual representations are recruited in mental rotation.

The distinction between MI and VWM abilities is further evidenced in the correlational analysis. Both TD children and children with ADHD demonstrated dissociated MI and VWM abilities in that there were no significant associations between each of the components of MI and VWM abilities in either group. However, it is important to note here that the sample size in of the ADHD group is underpowered to detect small effect sizes (see **Appendix A.1.1.** for sensitivity power analysis), therefore findings should be interpreted with caution. There were also no significant associations between symptoms of ADHD and either the components of MI or VWM measures in children with ADHD and TD children. The finding of a lack of a relationship between symptoms of ADHD and VWM is contrary to previous findings from mediation analysis that have shown an indirect effect of VWM performance on inattentive symptoms and hyperactivity/impulsivity symptoms in ADHD (Patros et al., 2015). Moreover, a longitudinal study demonstrated individual differences in greater improvements in VWM maintenance predicted reduced symptoms of ADHD (Karalunas et al., 2018). Given the relatively high VWM abilities evidenced in the sample of children with ADHD in this chapter, it is therefore perhaps not surprising that the maintenance and manipulation VWM measures were not associated with symptoms of ADHD in this group. In sum, contrary to hypotheses, children with ADHD present with typical patterns and age-appropriate levels of MI abilities as confirmed by both group-level and individual-level analyses.

The final experimental chapter of this thesis provides novel evidence to suggest that the subjective, sensory experience of MI is distinct from the visual precision and capacity of representations held in VWM. Previous research has demonstrated the established neural correlate of VWM, CDA, can index the visual precision at which information is held in mind and the number of visual items held in mind (Machizawa et al., 2012, 2020; Vogel & Machizawa, 2004). The findings presented in this thesis so far suggest a distinction between MI and VWM, which are interpreted to suggest that rather than this demonstrating wholly dissociated functions, it highlights individual differences in the recruitment of visual strategies in VWM. Alongside this, Pearson & Keogh's (2019) recent review argued that establishing how MI is recruited in VWM could provide a solution to competing theories for a role of frontal and visual areas in VWM and could explain individual differences in VWM capacity. Therefore, the study presented in **Chapter 4** was designed to establish exactly how individual differences in MI are related to neural and behavioural correlates of VWM. A classic, orientation-discrimination paradigm (Machizawa et al., 2012), where participants must identify whether a previously memorised item has been rotated clockwise or counter-clockwise, was adapted to include vividness and quantity ratings presented in both capacity-focused and quality-focused instruction blocks. The findings can be discussed in the context two important contributions. Firstly, it was evidenced that individuals exert wilful control over the precision and capacity of visual information held in mind during VWM and this varies dependent on whether they are following quality-focused vs. capacity-focused instructions and whether they expect to rate the vividness or quantity of their representations. Second, the findings suggest individuals have poor metacognitive insight into the content of their visual representations during VWM. The findings are considered in turn below.

With reference to the first novel contribution of this chapter, previous findings were extended by demonstrating that proportion correct and CDA were modulated by instruction and attended side. First, there was an indication of hemispheric specialisation of the visual regions, and this appeared to be dependent on instruction. Namely, proportion correct was greater in right attend trials compared to left attend trials in the coarse precision, quantity rating conditions only. Proportion correct was also greater in right attend compared to left attend trials in the in the 4-item condition of the capacity-focused, quantity rating block only. While the findings should be interpreted cautiously as they are based on trends for interactions, this is line with recent evidence for a distinguished role of the left lateral occipital cortex in VWM capacity (Machizawa et al., 2020). An attended side effect was also present in CDA but only in certain conditions. Specifically, greater CDA amplitudes were indexed in coarse compared to fine trials in the right attend trials of the capacity-focused blocks only. Modulations of CDA between coarse and fine trials was unexpected given that precision was not cued in this experiment. Moreover, the results from 3- and 4-way interactions in this study should be interpreted carefully due to limited number of trials and low power. Therefore, further investigation is required to warrant conclusions regarding the hemispheric specialisation of CDA.

A proactive control account of VWM is supported in that both proportion correct and CDA were modulated by instruction and the type of subjective MI rating. In line with prior hypotheses, proportion correct was greater in coarse precision compared to fine precision trials in the quality-focused instruction block only (as indicated by a trend towards a significant interaction between instruction and precision). This supports a proactive control account (Braver et al., 2007; Machizawa et al., 2012) in that it demonstrates that proportion correct was modulated when participants were following a quality-focused instruction, i.e., hold a precise mental image in mind. Interestingly, the expected interactions between ratings and instruction were not observed; vividness ratings were higher in smaller set sizes compared to larger set sizes and quantity ratings increased with increasing set size, however, contrary to expectations, and MI ratings did not differ between instruction blocks. This shows that while proportion correct is influenced by instruction, subjective MI ratings are not modulated in the same way.

Support for a proactive control account is further demonstrated in CDA amplitudes via a trend for an interaction between set size and rating type. This is vital to consider in that it suggests participants memory consumption (as indexed by CDA) was modulated depending on whether they were expecting to rate vividness ratings compared to quantity ratings. While trends are interpreted with caution, these findings support the notion that individuals exert wilful control over the visual precision and capacity of representations (Machizawa et al., 2012; Machizawa et al., 2020), and extend such findings by demonstrating that individuals not only exert this control following a specific cue but also when following blocked instructions and when expecting to provide qualitatively different (i.e., vividness vs. quantity) ratings of their representation.

In light of the findings that proportion correct and CDA were modulated depending on the type of rating participants expected, it might be anticipated that the hypotheses regarding the metacognitive link between MI and VWM would be supported. However, this was not the case. Trial-by-trial vividness and quantity ratings were included to examine individual differences in MI, following previous suggestions that individuals have good insight into their MI (Pearson et al., 2011). However, the results presented in **Chapter 4** suggest that individuals do not have good insight into the visual precision and capacity of representations in VWM. The results showed that while there was no difference in proportion correct between high vividness and low vividness trials, proportion correct was significantly greater in non-divergent compared to divergent quantity ratings. Furthermore, there was no difference in either CDA amplitude (indexing the content of visual representations held in mind during VWM maintenance) or ADAN amplitude (indexing the prioritisation of visual information in VWM) between high vividness and low vividness trials and non-divergent and divergent trials, respectively. This is of vital theoretical and methodological importance in that the findings demonstrate a disconnect between the subjective sensory experience of MI and the extent to which visual information is prioritised prior to the delay in VWM (ADAN) and held in mind during the delay in VWM (CDA). Overall, while individuals might have some insight into the number of items held in mind, the metacognitive link between MI and VWM is tentative at best.

## 5.3. Theoretical implications

The theoretical contributions of this thesis to current understanding can be summarised into two principal themes. The first is that this thesis provides convincing evidence that MI is not a unitary construct; it is a complex, multi-faceted function, which also differentiates in terms of implicit visual representations and the subjective sensory experience of MI. This has extensive implications for our current knowledge of MI, the development of MI and for methodological procedures moving forward. The second principal theoretical contribution is the demonstration that MI and VWM are in fact distinct from one another. This is demonstrated by comparing components of MI to components of VWM using both group-level and individual-level analyses (**Chapter 2** and **Chapter 3**) and by examining the metacognitive link between individual differences in the subjective sensory of experience of MI and VWM (**Chapter 4**). This ultimately suggests that the argument that MI and VWM are one and the same (Tong, 2013) is premature. The following section will explain each of the principal contributions in turn.

First, the findings highlight the importance of adopting a multi-faceted model of MI to fully examine MI abilities in both typical and atypical populations. Since the early theorisation of MI (Kosslyn, 1980; Kosslyn et al., 2006), a multi-component model has been largely neglected in favour of including one MI task. Such MI tasks tend to only instruct an individual to imagine a previously viewed stimulus (e.g., Dikjstra et al., 2017b; Lee et al., 2012; Stokes et al., 2009), which arguably only taps into the ability to generate an image. MI is inherently difficult to quantify in that there are extensive individual differences in the subjective sensory experience of MI (e.g., Dijkstra et al., 2019; Zeman et al., 2015). Therefore, the literature has thus far focused on using neuroimaging techniques to determine the format of representations recruited in MI and how mental images are constructed in adults (reviewed in detail in **Chapter 1, Section 1.1.**), and less on the separable components of MI. Evidence presented in this **Chapter 2** that demonstrates that the sub-components of MI develop separably throughout childhood and that some components become integrated in adulthood (image maintenance and mental rotation) highlights the importance of recruiting a separable-component model to fully elucidate the development of MI.

A separable-component model of MI in children was further supported by data-driven profile analyses presented in **Chapter 3** in the form of significant differences in performance on each of the components between most profiles. Prior to explaining the findings from the latent profile analysis, it is important to note some caveats. This method is relatively exploratory compared to hierarchical clustering methods in that no priors are set; the findings are data-driven with respect to modelling the probability of each case belonging to a specific profile (Goodman, 1974). This is therefore most appropriate for the current thesis in that there were no prior hypotheses with respect to the profile outcomes. Moreover, the sample size suggestion is 100 participants (Williams &

Kibwoski, 2016). The full sample in **Chapter 3** is just under this suggestion (88 participants), however the sample is of course imbalanced towards typically developing children. While this does not detract from the importance of considering individual differences in MI and VWM and the value of the findings, future research should recruit a representative sample of children with ADHD.

Three out of five profiles differentiated based on mental rotation performance (profiles 1, 2 and 4) and one profile (profile 3) presented with relatively low image generation and image maintenance abilities alongside relatively good mental rotation and image scanning. This specific distinction between more visual components (image generation and image maintenance) and the component involving spatial transformations (mental rotation) in children is important to consider. Although the research is in its infancy, recent investigations aiming to characterise the cognitive profile of those with Aphantasia sheds light on this distinction. Individuals with Aphantasia subjectively report intact spatial imagery, as measured by the Object-Spatial Imagery Questionnaire (OSIQ; Blajenkova et al., 2006), alongside absent visual imagery (Dawes et al., 2020; Keogh & Pearson, 2017). In addition, a recent study found mental rotation performance in those with Aphantasia did not differ from controls without Aphantasia (Pounder et al., 2018). Thus, this is in line with a distinction between visual components of MI and mental rotation found in children in this thesis.

Comparatively, there is contention in the literature as to whether mental rotation involves the maintenance of visual information. On the one hand, it has been argued that while visual information is encoded in mental rotation tasks, only the orientation-dependent spatial information is extracted to successfully complete the mental rotation (Liesefeld & Zimmer, 2013). On the other hand, studies have shown ERP components are elicited during visual short-term memory (i.e., only one visual item and no requirement for manipulation; VSTM) are also elicited during mental rotation (Prime & Jolicoeur, 2010). Moreover, the rotation-related negativity (RRN) component observed in posterior electrodes has been found to increase in amplitude as a function of increasing degree of rotation (Riečanský et al., 2013; Riečanský & Jagla, 2008). Given that both are slow posterior negative ERPs, it has been interpreted that mental rotation involves the rotation of a visual image. However, it has also been shown that the negative slow wave observed in VSTM is not significantly related to the

RRN component, which was taken to suggest that mental rotation does not involve the retention of a visual image (Riečanský et al., 2013). Taken in the context of the evidence presented in this thesis, this is perhaps not surprising. A likely explanation is that there are individual differences in the flexible recruitment of visual information during mental rotation. The evidence presented in this thesis offers important insights into this relationship. Firstly, evidence for a relationship between image maintenance and mental rotation in adults demonstrates that mental rotation can include the maintenance of visual information. Secondly, there was no relationship between the mental rotation and image maintenance at the group-level in children. Moreover, data-driven profile analysis demonstrated variable profiles of mental rotation and image maintenance ability in that profiles were characterised as both good mental rotation and poor image maintenance (profile 3) and comparatively good mental rotation and good image maintenance (profiles 4 and 5). Thus, it would be too crude a conclusion to determine that mental rotation does not recruit visual information, and instead the evidence points towards individual differences in the recruitment of visual information in mental rotation.

The evidence for a dissociation between MI and VWM in this thesis is contrary to the current narrative in the literature and forms a vital contribution to the debate. Distinct abilities were first demonstrated by comparing performance on components of MI and VWM measures in adults, TD children and children with ADHD (**Chapter 2 and 3**). The findings reported in **Chapter 2** and **Chapter 3** demonstrated the MI and VWM are not synonymous, and it was interpreted that the recruitment of visual mental images in VWM is likely dependent on individual differences. This was then examined directly in **Chapter 4**. The findings presented in **Chapter 4** importantly demonstrate that individuals have flexible control over the visual precision and capacity of visual representations. Importantly, this was modulated dependent on whether participants were following quality-focused or quantity-focused instructions and whether participants were expecting to rate vividness or quantity ratings. Without the inclusion of subjective MI ratings in this experiment, an MI researcher might conclude this reflects the flexible recruitment of MI strategies in VWM. However, examining the metacognitive MI ratings revealed that these were disconnected from the neural correlates of visual precision and capacity in VWM. As outlined in **Chapter 4**, a likely explanation for the seemingly contradictory findings could be the nature in which MI and VWM are assessed. In **Chapters 2** and **3**, MI and VWM are assessed separately, tasks were counterbalanced for each participant and no prior instructions regarding visual or MI strategies were given in the VWM tasks. In the previous fMRI study demonstrating shared visual representations between the two functions, MI and VWM were presented in the same task separated by a MI cue and a VWM cue, respectively (Albers et al., 2013), therefore participants expected to be required to recruit MI within the task. Similarly, in the VWM task presented in **Chapter 4**, participants were instructed to hold either a precise mental image in mind or hold as many visual items in mind as possible and they also expected to provide subjective MI ratings. Hence, evidence for wilful control of the precision and capacity of visual representations was found and this finding does not contradict the assertion that visual representations are recruited in both MI and VWM. On the whole, the findings taken together demonstrate that visual representations of varying degrees of precision and capacity are implicated in VWM (as demonstrated in **Chapter 4**), however, MI or visual strategies are not necessarily a requirement of VWM (as evidenced in **Chapters 2** and **3**).

Another distinction is highlighted with respect to the metacognitive link between the subjective experience of MI and behavioural and neural correlates of VWM. This Is seemingly contradictory to evidence suggesting the subjective sensory experience of MI is associated with selective activation of the early visual areas (e.g., Cui et al., 2007; Lee et al., 2012; Dijkstra et al., 2017b). However, this has been previously demonstrated with fMRI as opposed to the use of EEG in the current study. Therefore, while the previous studies measure selective activation of early visual areas, EEG in this study measures event-related potentials with millisecond precision, thus is it not expected that these findings would align completely. The finding that the subjective experience of MI is distinct from the behavioural and neural correlates of VWM is not only important theoretically, but it highlights an important methodological distinction in measurement. Preliminary studies investigating the cognitive profile of Aphantasia largely recruit subjective MI questionnaires (Dawes et al., 2020; A. Zeman et al., 2015). The findings presented in this thesis therefore raise the question as to whether Aphantasia is a specific deficit in the subjective sensory experience of MI as opposed to the recruitment and maintenance of visual representations to support memory. Research is currently limited, but this might explain why a case study of an individual with Aphantasia found that they performed at the expected level in a VWM task (Jacobs et al., 2018). Taken together, the findings provide novel theoretical and methodological contributions with respect to the relationship between MI and VWM.

## 5.4. Practical implications

The implications of this thesis extend beyond the theoretical contributions to practical implications. Findings for separable components and distinct functions of MI and VWM have implications for how MI supports learning. Previous studies have indicated a positive relationship between MI and academic outcomes. MI has been found to predict geometric learning in children age 10-12 years (Bizzaro et al., 2018) and MI skills at age 4-5 years were found to longitudinally predict mathematics skills (measured with a standardised mathematics test including calculation, enumeration, and numerical facts), reading and writing proficiency at age 6-7 years (Guarnera et al., 2019). However, different tasks were recruited to measure MI in each study, and these did not always conform to a separable component model of MI. Bizzaro et al.'s (2018) study adopts one MI measure derived from an imagery and spatial skills battery (Mammarella et al., 2012). This measure comprised items requiring composing and decomposing embedded figures, which arguably measures more general non-verbal reasoning as opposed to the ability to generate, maintain, or manipulate mental images. Guarnera et al.'s (2019) study adopts a component model of MI however the exact components being measured and therefore the approach to the task is unclear. For example, in the image generation measure, participants learn object-letter pairs (e.g., picture of a home paired with the letter "A") and are then required to identify the object from the presentation of the paired letter. However, this task could be solved using a verbalising strategy, i.e., verbally recalling the home is paired with "A". In this sense, it is not surprising that performance on this task predicted writing abilities. Clearly, research is required to establish how individual differences in the separable-component model of MI contributes to academic outcomes throughout childhood. A central aim of this thesis was to provide a novel battery of tasks that are sensitive to individual differences in abilities on each component of MI in children and adults. This battery of tasks was found to be effective in characterising MI abilities from age 6 years in typical development and in children with ADHD, which therefore provides a useful tool for future research investigating how MI supports learning in both typical and atypical development.

The battery of tasks introduced in this thesis has since been implemented in a study[1] examining how components of MI contribute to mathematical skills in children (Bates et al., 2021). Research has shown that spatial skills are positively associated with mathematical skills (Geer et al., 2019; Gilligan et al., 2017), with some demonstrating a causal link between the two. Specifically, training mental rotation has been found to lead to gains in mathematical calculation ability (Gilligan et al., 2019; Mix et al., 2020). Spatial visualisation is often used as an explanation of the mechanistic link between spatial skills, such as mental rotation, and mathematical skills, however the definition of spatial visualisation is unclear. Some definitions stress the importance of visual representations, as well as spatial transformations (Linn & Petersen, 1985), whereas others refer only to transformation of spatial properties (Lowrie et al., 2019). Adopting a separable-component model of MI in this context allows for the investigation of how visual components (image generation, image maintenance, image scanning) and components that involve spatial transformations (mental rotation) contribute to mathematical skills to provide clarity on the underlying MI mechanisms. In this study it was found that only mental rotation significantly predicted mathematical calculation skills over and above age. This provides important insights into the mechanisms underpinning mathematical calculation skills in that it suggests that spatial transformations are important as opposed to the visual precision of representations (Bates et al., 2021). This therefore has implications regarding the types of strategies that might be encouraged or instructions that might be given in the mathematics classroom. Current research is limited to mathematical calculation skills, however this battery of MI tasks could be applied to future research to examine other types of mathematics skills, problem-solving skills for instance, to establish the role of MI in mathematical learning.

---

[1] Data from Chapter 2 was used in the study presented in Bates et al. (2021), however this study is not reported in this thesis

In addition, the identification of typical mental rotation abilities in children with ADHD has implications for intervention. The evidence presented here suggests that children with ADHD perform at the typical level in mental rotation. This contrasts with previous suggestions of impairments compared to TD controls (e.g., Silk et al., 2005; Vance et al., 2007). However, it is important to note that in the previous studies, as well as in the study presented in **Chapter 3**, sample sizes of children with ADHD were very small. Therefore, while tentative conclusions can be made here, further research is required to assess the full range of mental rotation abilities in ADHD. Taken in the context of research demonstrating the importance of mental rotation in mathematics abilities, evidence for age-appropriate mental rotation skills in children with ADHD reported in this thesis is promising. Poorer academic and later life outcomes compared to typical peers are commonly identified in children with ADHD (Best et al., 2011). Despite the identification that poor VWM contributes to poor academic achievement in ADHD, particularly in mathematics (Bull et al., 2008; Friedman et al., 2018), the evidence suggests that VWM training does not lead to gains in either VWM or mathematics (Cortese et al., 2015; Rapport et al., 2013). Therefore, if training in mental rotation can lead to gains in mathematical calculation skills in TD children (Gilligan et al., 2019; Mix et al., 2020), it might be possible to improve mathematical skills in children with ADHD to compensate for possible deficits due to poor VWM by training mental rotation.

Evidence for typical performance in MI and VWM in children with ADHD demonstrated in this thesis directly pertains to the argument for adopting a spectrum approach to investigating neurodevelopmental disorders, which has implications for understanding cognition in ADHD moving forward. Investigation into the neuropsychological profile of ADHD has largely focused on deficits and impairments (e.g., Barkley, 1997; Castellanos et al., 2006; Zelazo & Carlson, 2012). However, the recent recruitment of data-driven analyses has revealed extensive individual differences in a range of cognitive functions in both typical development and in ADHD (Fair et al., 2011; Dajani et al., 2014; Campez et al., 2020). This has contributed to an emerging argument for shifting from a deficit-focus in ADHD towards considering the resources and skills individuals with ADHD have and can develop (see Lesch, 2018 for discussion). Moreover, it has recently been argued that theoretical accounts of neurodevelopmental disorders that do not capture heterogeneity are flawed (Astle & Fletcher-Watson, 2020).

The evidence presented in this thesis directly contributes to this debate in that evidence is provided for typical patterns of performance not only in mental rotation but more broadly in each of the components of MI in ADHD. This was also supported in data-driven analyses that demonstrated individual differences were not ADHD-specific. Taken together, this thesis provides convincing evidence for heterogeneity in the cognitive profiles of ADHD with respect to MI and VWM and demonstrates the importance of data-driven approaches to analysis.

## 5.5. Limitations and future research

It is firstly important to acknowledge that the developmental comparisons in this thesis are cross-sectional rather than longitudinal and are limited to children of primary school age and adults. Therefore, further research is required to examine how MI develops throughout the adolescent years through to adulthood and from adulthood through to older adults. While the oldest developmental age group (10- to 11-year-olds) tested in this thesis appeared to perform at adult-like levels in image generation and image scanning, accuracy in image maintenance and mental rotation appeared to be developing beyond 10- to 11-year-olds. To my knowledge, only one other study has tested components of MI in an adolescent group (age 14 years; Kosslyn et al., 1990) and adult-like abilities were indicated. However, research is required to address this gap in knowledge with respect to the adolescent years. Moreover, the conclusion that adult-like abilities have been reached should be tentative given the age range of the adult group. Cross-sectional analyses on individuals aged 5-80 years have revealed extensive development of VWM capacity up to around 30 years of age (Alloway & Alloway, 2013). The mean age of adults tested in **Chapter 2** was 26 years old, therefore there could be further maturity of MI beyond this group. In older adults (mean age of 63), decline in the image generation, image maintenance and mental rotation abilities have been noted alongside comparable performance in image scanning to younger adults (mean age of 20 years) (Dror & Kosslyn, 1994). In addition, a recent VWM training study revealed an instruction visualisation strategy was less useful for older adults (mean age of 69) than for younger adults (mean age of 22) (Forsberg et al., 2020). Taken together, this demonstrates that to establish the full picture of individual differences in MI and VWM, a lifespan approach is required. While tentative conclusions can be made regarding developmental progression up to age 11 years, it is not possible to fully determine the developmental progression of the components of MI based on the data presented in this thesis.

It should also be noted that planned analyses stated correlation analyses would be conducted for each developmental age group; 6- to 7-year-olds, 8- to 9-year-olds and 10- to 11-year-olds, respectively, however, sensitivity power analyses revealed the study was not powered for separate correlation analyses. Correlations at each age range would be particularly useful given that previous research has suggested relationships between the components of MI might differ dependent on age. In the initial study, Kosslyn et al. (1990) found no associations between any of the components of MI, except for between mental rotation and image scanning in 8-year-olds. In contrast, Wimmer et al. (2017) found no associations between mental rotation and image scanning when controlling for age, however, when age groups were assessed separately, significant associations were found between mental rotation and image scanning RTs in 6-, 10-year-olds and adults only. Future studies powered to detect small effect sizes in correlation analyses are required to establish how components of MI are related to one another throughout childhood.

The samples included in **Chapter 2** and **3** are limited with the respect to the spectrum of VWM abilities and the sample in **Chapter 3** is limited with respect to sample size. As outlined in **Chapter 3**, the ADHD group correlation analyses are underpowered due to a smaller final sample size than outlined in the preregistration. This was because data collection had to be terminated early due to extended COVID-19 restrictions and school closures. Findings regarding associations between components of MI and VWM should therefore be interpreted with caution. This also reflects the wider issue in neurodevelopmental disorder research of small and selective sample sizes, partly due to time and resource constraints, which therefore limits conclusions regarding neuropsychological profiles (Astle & Fletcher-Watson, 2020). It is also notable that the sample of children with ADHD, when compared to the literature, seem to be performing at the higher end of the spectrum of VWM abilities. Previous literature has repeatedly identified impairments in VWM in children with ADHD compared to TD controls (Kasper et al., 2012), therefore the sample in **Chapter 3** appears to only include a cross-section of children with ADHD with typical abilities in VWM. Therefore, conclusions regarding the relationship between components of MI and VWM in children with ADHD made here are limited to the sample tested and further investigation with a representative sample of children with ADHD is required.

A notable limitation of **Chapter 4** is the number of trials in the analysis of individual differences in metacognitive ratings. Because the number of trials in the individual difference analyses (i.e., comparisons of high and low vividness ratings and non-divergent and divergent quantity ratings) was dependent on the participant response, in some cases there were no trials to analyse and this also meant that variables collapsed across conditions (e.g., low vividness ratings collapsed across instruction, set size, precision and attended side) were comprised of different numbers of trials per participant. Given the variability in responses, it could be that the low vividness dependent variable might be limited to higher set sizes (i.e., some participants didn't rate any 4-item trials as high vividness) and that the high vividness dependent variable might be limited to smaller set sizes. As this was the first study to examine how subjective MI ratings of vividness and quantity relate to behavioural and neural outcomes, it was not possible anticipate this level of variability. Therefore, the findings based on variables collapsed across conditions should be interpreted cautiously and a metacognitive link between MI and VWM cannot be definitively ruled out based on this data. To confirm the findings suggesting a dissociation between subjective MI ratings and VWM neural correlates, future research should sample participants dependent on their rated vividness and quantity, respectively, to ensure appropriate trial numbers per condition.

The findings from **Chapter 4** also provide promising avenues for future research. To examine both the role of visual and frontal regions in the relationship between MI and VWM, the CDA component observed in the posterior electrodes was analysed and the ADAN component observed in the frontal electrodes was analysed. The choice of components was also important in terms of timing and the application of EEG allowed for the distinction between components early in the VWM trial compared to later in the VWM trial. Namely, the ADAN component is measured following the cue onset at the start of the trial and the CDA component is measured during the delay period following sample onset. However, the spatial resolution of EEG is restricted, and the EEG analyses presented in **Chapter 4** can only give a crude indication of the relative contribution of both frontal and visual regions. One previous study has applied a similar logic using magnetoencephalography (MEG) which allows spatially and temporally precise measurements. Here it was demonstrated that early activity in the frontal regions propagates to later activity in the visual areas during MI (Dijkstra et al., 2020). Therefore, it would be informative to conduct a study examining the relationship between MI and VWM using MEG to further investigate the role of frontal and visual regions.

Finally, the contributions of this thesis are limited to the visual domain. While the investigation of working memory is largely investigated in isolated domains, with the most extensively researched being VWM, the real-world environments in which we learn, and process information are multisensory in nature. Crucially, the brain readily integrates multisensory information at early (<100ms) processing stages in primary sensory cortices, prior to attention modulations (De Meo et al., 2015; Murray et al., 2016). Evidence has demonstrated that pairing a sound with a visual target improves visual search (Van der Burg et al., 2008) and enhances visual attentional capture via bottom-up control (Folk et al., 1992; Matusz & Eimer, 2011). Thus, the simultaneous presentation of audiovisual information increases the saliency of visual stimuli and draws attention to it more strongly. The impact of how audiovisual saliency modulates behavioural and neural correlates of VWM is currently unknown. However, research has investigated how object memory is influenced by encoding audiovisual information. This has been found to be dependent on semantic congruency: when audiovisual information is semantically congruent, e.g., the image of a dog paired with the sound of a bark, subsequent visual object memory is improved compared to stimuli that were experienced only visually (Matusz et al., 2017; Murray et al., 2004). Notably, when audiovisual information lacks a semantic relationship, e.g., the image of a dog is paired with a simple beep, the memory benefits show strong individual variability, with some individuals showing improved object memory whereas others are impaired, compared to encoding a visual-only stimulus (Thelen et al., 2014). To establish how MI might support memory and learning in real-world environments, such as the classroom, it would be valuable to build on the evidence presented in this thesis in the visual domain to examine how the encoding of audiovisual information contributes to individual differences in MI and VWM.

## 5.6. Concluding remarks

The understanding of the development of MI has been incomplete to date, with limited investigation into the components of MI and into the extent to which a depictive theory of MI is supported in childhood. This thesis provides novel evidence to provide support for a depictive theory of MI in development in that children from age 6 were able to generate, maintain, manipulate, and shift attention across varying distances in visual mental images. Evidence for a separable-component model of MI is supported in childhood alongside the suggestion that image maintenance and mental rotation abilities become more integrated in adulthood. This has important theoretical contributions to the debate as to whether mental rotation involves the recruitment of visual information. Specifically, it appears that mental rotation can involve visual information and this likely depends on individual differences and the developmental stage. In sum, the evidence presented in this thesis demonstrates that MI is a multi-faceted function, rather than a unitary construct.

While extensive research has investigated VWM abilities in children with ADHD, to my knowledge, no previous study has investigated MI abilities in ADHD. This thesis characterised abilities in components of MI alongside VWM in children with ADHD for the first time. Between-group comparisons revealed children with ADHD presented with typical patterns of performance in each component of MI alongside age-appropriate levels of ability. This was supported in data-driven analyses which showed individual differences in MI and VWM were not specific to ADHD. This has both theoretical and practical implications. Firstly, it further demonstrates support for a separable-component model of MI and extends these findings to the ADHD population. Moreover, this is an important contribution to the movement in neurodevelopmental research arguing for the importance of going beyond core-deficit approaches and considering individual differences. From a practical perspective, this could provide an opportunity to engage children with ADHD in classroom learning. Research in typically developing populations has demonstrated causal links between mental rotation and mathematical skills. In this thesis, children with ADHD were found to present with typical mental rotation abilities. Future research should explore the relationship between mental rotation and mathematical skills in children with ADHD, which could in turn lead to training studies. Importantly, this thesis provides a framework for examining abilities in the components of MI in children from age 6 years suitable for both typical and atypical populations, which will pave the way for investigating how MI contributes to classroom learning.

Current and limited evidence has led to the argument that MI and VWM are in fact synonymous. The evidence presented in this thesis contests this view. The relationship between components of MI and VWM was tested directly for the first time in adults, TD children and children with ADHD. Evidence for a dissociation between MI and VWM abilities was found in each of the populations, and this was interpreted to suggest that the recruitment of visual strategies in VWM might be dependent on individual differences in MI. This was directly examined in the final experimental chapter. Importantly, it was demonstrated that individuals can exert wilful control over the visual precision and capacity of visual representations depending on the instruction being followed and the type of subjective MI rating they expect to report. Contrary to expectations, it was also found that there was a disconnect between the subjective experience of MI and the visual precision and capacity of visual representations in VWM. Thus, there is a distinction between the subjective experience of MI and the implicit visual representations recruited in memory. This is vital moving forward as it suggests that while the subjective sensory experience of MI might be variable, for example in those with Aphantasia, individuals still have flexible control over the precision and capacity of visual representations that support memory. This contribution of evidence in the visual domain in this thesis should be extended in future research to examine how multisensory information is encoded and how this in turn impacts memory in real-world environments.

# References

Adam, K. C. S., Robison, M. K., & Vogel, E. K. (2018). Contralateral delay activity tracks fluctuations in working memory performance. *Journal of Cognitive Neuroscience*, *30*(9), 1229–1240. https://doi.org/10.1162/jocn_a_01233

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Alderson, R. M., Rapport, M. D., Hudec, K. L., Sarver, D. E., & Kofler, M. J. (2010). Competing core processes in attention-deficit/hyperactivity disorder (ADHD): Do working memory deficiencies underlie behavioral inhibition deficits? *Journal of Abnormal Child Psychology*, *38*(4), 497–507. https://doi.org/10.1007/s10802-010-9387-0

Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-Deficit / Hyperactivity Disorder and Behavioral Inhibition : A Meta-Analytic Review of the Stop-signal Paradigm. *Journal of Abnormal Child Psychology*, *30*(5), 745–758. https://doi.org/10.1007/s10802-007-9131-6

Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology*, *25*(1), 84–93. https://doi.org/10.1080/20445911.2012.748027

Alloway, T. P., Elliott, J., & Place, M. (2010). Investigating the Relationship Between Attention and Working Memory in Clinical and Community Samples. *Child Neuropsychology*, *16*(3), 242–254. https://doi.org/10.1080/09297040903559655

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and Visuospatial Short-Term and Working Memory in Children: Are They Separable? *Child Development*, *77*(6), 1698–1716. https://doi.org/10.1111/j.1467-8624.2006.00968.x

Astle, D. E., & Fletcher-Watson, S. (2020). Beyond the Core-Deficit Hypothesis in Developmental Disorders. *Current Directions in Psychological Science*, *29*(5), 431–437. https://doi.org/10.1177/0963721420925518

Astle, D., Holmes, J., Kievit, R., & Gathercole, S. E. (2021). *The transdiagnostic revolution in neurodevelopmental disorders*.

https://doi.org/10.31234/OSF.IO/4G3A2
Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews: Neuroscience*, *4*(October), 829–839. https://doi.org/10.1038/nrn1201
Baddeley, A. D., & Andrade, J. (2000). Working Memory and the Vividness of Imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126
Bae, G., & Luck, S. J. (2019). What happens to an individual visual working memory representation when it is interrupted? *British Journal of Psychology*, *110*(2), 268–287. https://doi.org/10.1111/bjop.12339
Barkley, R. A. (1997). Behavioral Inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94.
Bates, K. E., Gilligan-Lee, K., & Farran, E. K. (2020). Reimagining mathematics: The role of mental imagery as a predictor of mathematical calculation skills in childhood. *PsyArxiv*. https://doi.org/10.31234/osf.io/r7gnc
Bates, K. E., Gilligan-Lee, K., & Farran, E. K. (2021). Reimagining Mathematics: The Role of Mental Imagery in Explaining Mathematical Calculation Skills in Childhood. *Mind, Brain, and Education*. https://doi.org/10.1111/mbe.12281
Behrmann, M., Winocur, G., & Moscovitch, M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature*, *359*(6396), 636–637. https://doi.org/10.1038/359636a0
Berg, D. H. (2008). Working memory and arithmetic calculation in children: The contributory roles of processing speed, short-term memory, and reading. *Journal of Experimental Child Psychology*, *99*(4), 288–308. https://doi.org/10.1016/j.jecp.2007.12.002
Bergwerff, C. E., Luman, M., Weeda, W. D., & Oosterlaan, J. (2019). Neurocognitive Profiles in Children With ADHD and Their Predictive Value for Functional Outcomes. *Journal of Attention Disorders*, *23*(13), 1567–1577. https://doi.org/10.1177/1087054716688533
Berryhill, M. E., Phuong, L., Picasso, L., Cabeza, R., & Olson, I. R. (2007). Parietal lobe and episodic memory: bilateral damage causes impaired free recall of autobiographical memory. *The Journal of Neuroscience*, *27*(52), 14415–14423. https://doi.org/10.1523/JNEUROSCI.4163-07.2007
Berwid, O. G., Curko Kera, E. A., Marks, D. J., Santra, A., Bender, H. A., &

Halperin, J. M. (2005). Sustained attention and response inhibition in young children at risk for Attention Deficit/Hyperactivity Disorder. *Journal of Child Psychology and Psychiatry*, *46*(11), 1219–1229. https://doi.org/10.1111/j.1469-7610.2005.00417.x

Best, J. R., Miller, M. H., & Naglieri, J. A. (2011). Relations between Executive Function and Academic Achievement from Ages 5 to 17 in a Large, Representative National Sample. *Learning and Individual Differences*, *21*(4), 327–336. https://doi.org/10.1016/j.lindif.2011.01.007.Relations

Bied, A., Biederman, J., & Faraone, S. (2017). Parent-based diagnosis of ADHD is as accurate as a teacher-based diagnosis of ADHD. *Postgraduate Medicine*, *129*(3), 375–381. https://doi.org/10.1080/00325481.2017.1288064

Bizzaro, M., Giofrè, D., Girelli, L., & Cornoldi, C. (2018). Arithmetic, working memory, and visuospatial imagery abilities in children with poor geometric learning. *Learning and Individual Differences*, *62*(January), 79–88. https://doi.org/10.1016/j.lindif.2018.01.013

Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. *Applied Cognitive Psychology*, *20*(2), 239–263. https://doi.org/10.1002/acp.1182

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Boccia, M., Sulpizio, V., Teghil, A., Palermo, L., Piccardi, L., Galati, G., & Guariglia, C. (2019). The dynamic contribution of the high-level visual cortex to imagery and perception. *Human Brain Mapping*, *40*(8), 2449–2463. https://doi.org/10.1002/hbm.24535

Borst, G., Kievit, R., Thompson, W., & Kosslyn, S. (2011). Mental rotation is not easily cognitively penetrable. *Journal of Cognitive Psychology*, *23*(1), 60–75. https://doi.org/10.1080/20445911.2011.454498

Borst, G., & Kosslyn, S. M. (2008). Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory & Cognition*, *36*(4), 849–862. https://doi.org/10.3758/MC.36.4.849

Borst, G., & Kosslyn, S. M. (2010). Individual differences in spatial mental imagery. *Quarterly Journal of Experimental Psychology*, *63*(10), 2031–2050. https://doi.org/10.1080/17470211003802459

Borst, G., Kosslyn, S. M., & Denis, M. (2006). Different cognitive processes in two image-scanning paradigms. *Memory and Cognition*, *34*(3), 475–490. https://doi.org/10.3758/BF03193572

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In *Variation in Working Memory* (pp. 76–106). Oxford University Press.

Breedlove, J. L., St-Yves, G., Olman, C. A., & Naselaris, T. (2020). Generative Feedback Explains Distinct Brain Activity Codes for Seen and Mental Images. *Current Biology*, *30*(12), 2211-2224.e6. https://doi.org/10.1016/j.cub.2020.04.014

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228. https://doi.org/10.1080/87565640801982312

Burnett Heyes, S., Zokaei, N., van der Staaij, I., Bays, P. M., & Husain, M. (2012). Development of visual working memory precision in childhood. *Developmental Science*, *15*(4), 528–539. https://doi.org/10.1111/j.1467-7687.2012.01148.x

Campez, M., Raiker, J. S., Sarver, D. E., Friedman, L. M., Orban, S. A., & Rapport, M. D. (2020). Working Memory Capacity and ADHD Symptoms in Boys: Examining the Heterogeneity of Working Memory Functioning Using Latent Profile Analysis. *Journal of Psychopathology and Behavioral Assessment*, *42*(3), 450–463. https://doi.org/10.1007/s10862-019-09762-x

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*(3), 117–124. https://doi.org/10.1016/j.tics.2006.01.011

Chatterjee, A., & Southwood, M. H. (1995). Cortical blindness and visual imagery. *Neurology*, *45*(12), 2189–2195. https://doi.org/10.1212/WNL.45.12.2189

Chen, W., Kato, T., Zhu, X. H., Ogawa, S., Tank, D. W., & Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *NeuroReport*, *9*(16), 3669–3674.

https://doi.org/10.1097/00001756-199811160-00019

Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012). Imagery and Perception Share Cortical Representations of Content and Location. *Cerebral Cortex*, *22*(2), 372–380. https://doi.org/10.1093/cercor/bhr106

Cochrane, B. A., Townsend, B., Lee, R. C., Legere, J. K., Milliken, B., & Shedden, J. M. (2020). Visual imagery influences attentional guidance during visual search: Behavioral and electrophysiological evidence. *Attention, Perception, and Psychophysics*, *83*(1), 58–66. https://doi.org/10.3758/s13414-020-02208-7

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. In *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Commodari, E., Guarnera, M., Di Stefano, A., & Di Nuovo, S. (2020). Children Learn to Read: How Visual Analysis and Mental Imagery Contribute to the Reading Performances at Different Stages of Reading Acquisition. *Journal of Psycholinguistic Research*, *49*(1), 59–72. https://doi.org/10.1007/s10936-019-09671-w

Conners, C. K. (2008). *Conners third edition (Conners 3)*. Western Psychological Services.

Constantinidis, C., Franowicz, M. N., & Goldman-Rakic, P. S. (2001). The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature Neuroscience*, *4*(3), 311–316. https://doi.org/10.1038/85179

Corbetta, M. (1993). Positron emission tomography as a tool to study human vision and attention. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(23), 10901–10903. https://doi.org/10.1073/pnas.90.23.10901

Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., Holtmann, M., Santosh, P., Stevenson, J., Stringaris, A., Zuddas, A., & Sonuga-Barke, E. J. S. (2015). Cognitive training for attention-deficit/hyperactivity disorder: Meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(3), 164–174. https://doi.org/10.1016/j.jaac.2014.12.010

Cortina, J. M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, *78*(1), 98–104.

https://doi.org/10.1037/0021-9010.78.1.98

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia*, *49*(6), 1401–1406. https://doi.org/10.1016/j.neuropsychologia.2011.01.035

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. In *Educational Psychology Review* (Vol. 26, Issue 2, pp. 197–223). Springer New York LLC. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N. (2016). Working Memory Maturation: Can We Get at the Essence of Cognitive Growth? *Perspectives on Psychological Science*, *11*(2), 239–264. https://doi.org/10.1177/1745691615621279

Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, *19*(2), 261–270. https://doi.org/10.1016/S1053-8119(03)00049-1

Cragg, L., Richardson, S., Hubber, P. J., Keeble, S., & Gilmore, C. (2017). When is working memory important for arithmetic? The impact of strategy and age. *PLoS ONE*, *12*(12), 1–18. https://doi.org/10.1371/journal.pone.0188693

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vision Research*, *47*(4), 474–478. https://doi.org/10.1016/j.visres.2006.11.013

Dajani, D. R., Llabre, M. M., Nebel, M. B., Mostofsky, S. H., & Uddin, L. Q. (2016). Heterogeneity of executive functions among comorbid neurodevelopmental disorders. *Scientific Reports*, *6*, 1–10. https://doi.org/10.1038/srep36566

Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-65705-7

De Meo, R., Murray, M. M., Clarke, S., & Matusz, P. J. (2015). Top-down control and early multisensory processes: Chicken vs. egg. *Frontiers in Integrative Neuroscience*, *9*(March), 1–6. https://doi.org/10.3389/fnint.2015.00017

Dean, A. L., & Harvey, W. O. (1979). An information-processing analysis of a Piagetian imagery task. *Developmental Psychology*, *15*(4), 474–475. https://doi.org/10.1037/0012-1649.15.4.474

Dentico, D., Cheung, B. L., Chang, J. Y., Guokas, J., Boly, M., Tononi, G., & Van Veen, B. (2014). Reversal of cortical information flow during visual imagery as compared to visual perception. *NeuroImage*, *100*, 237–243. https://doi.org/10.1016/j.neuroimage.2014.05.081

Dijkstra, N., Ambrogioni, L., Vidaurre, D., & van Gerven, M. A. J. (2020). Neural dynamics of perceptual inference and its reversal during imagery. *ELife*, *9*, 1–19. https://doi.org/10.7554/eLife.53588

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2017a). Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *The Journal of Neuroscience*, *37*(5), 1367–1373. https://doi.org/10.1523/jneurosci.3022-16.2016

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences*, *23*(5), 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Dijkstra, N., Hinne, M., Bosch, S. E., & van Gerven, M. A. J. (2019). Between-subject variability in the influence of mental imagery on conscious perception. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-52072-1

Dijkstra, N., Mostert, P., de Lange, F. P., Bosch, S., & van Gerven, M. A. J. (2018). Differential temporal dynamics during visual imagery and perception. *ELife*, *7*, 1–16. https://doi.org/10.7554/eLife.33904

Dijkstra, N., Zeidman, P., Ondobaka, S., Van Gerven, M. A. J., & Friston, K. (2017b). Distinct Top-down and Bottom-up Brain Connectivity during Visual Perception and Imagery. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-05888-8

Donolato, E., Giofrè, D., & Mammarella, I. C. (2017). Differences in verbal and visuospatial forward and backward order recall: A review of the literature. *Frontiers in Psychology*, *8:663*. https://doi.org/10.3389/fpsyg.2017.00663

Drisdelle, B. L., Aubin, S., & Jolicoeur, P. (2017). Dealing with ocular artifacts on lateralized ERPs in studies of visual-spatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology*, *54*(1), 83–99. https://doi.org/10.1111/psyp.12675

Dror, I. E., & Kosslyn, S. M. (1994). Mental imagery and aging. *Psychology and Aging*, *9*(1), 90–102. https://doi.org/10.1037/0882-7974.9.1.90

Estes, D. (1998). Young Children’s Awareness of Their Mental Activity: The Case of Mental Rotation. *Child Development*, *69*(5), 1345. https://doi.org/10.2307/1132270

Fair, D. A., Bathula, D., Nikolas, M., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences*, *109*(17), 6769–6774. https://doi.org/10.1073/pnas.1115365109

Farran, E. K., Purser, H. R. M., Courbois, Y., Ballé, M., Sockeel, P., Mellier, D., & Blades, M. (2015). Route knowledge and configural knowledge in typical and atypical development: A comparison of sparse and rich environments. *Journal of Neurodevelopmental Disorders*, *7*(1), 37. https://doi.org/10.1186/s11689-015-9133-6

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47. https://doi.org/10.1093/cercor/1.1.1

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications Inc.

Finke, R. A., & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(2), 142–147. https://doi.org/10.1037/0278-7393.8.2.142

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary Covert Orienting Is Contingent on Attentional Control Settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030–1044. https://doi.org/10.1037/0096-1523.18.4.1030

Forsberg, A., Fellman, D., Laine, M., Johnson, W., & Logie, R. H. (2020). Strategy mediation in working memory training in younger and older adults.

*Quarterly Journal of Experimental Psychology*, *73*(8), 1206–1226. https://doi.org/10.1177/1747021820915107

Frick, A., Ferrara, K., & Newcombe, N. S. (2013). Using a touch screen paradigm to assess the development of mental rotation between 3 and 5 years of age. *Cognitive Processing*, *14*(2), 117–127. https://doi.org/10.1007/s10339-012-0534-0

Frick, A., Hansen, M. A., & Newcombe, N. S. (2013). Development of mental rotation in 3- to 5-year-old children. *Cognitive Development*, *28*(4), 386–399. https://doi.org/10.1016/j.cogdev.2013.06.002

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends in Cognitive Sciences*, *18*(10), 536–542. https://doi.org/10.1016/j.tics.2014.05.011

Friedman, L. M., Rapport, M. D., Orban, S. A., Eckrich, S. J., & Calub, C. A. (2018). Applied Problem Solving in Children with ADHD: The Mediating Roles of Working Memory and Mathematical Calculation. *Journal of Abnormal Child Psychology*, *46*(3), 491–504. https://doi.org/10.1007/s10802-017-0312-7

Fukusima, S. S., Loomis, J. M., & Da Silva, J. A. (1997). Visual Perception of Egocentric Distance as Assessed by Triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(1), 86–100. https://doi.org/10.1037/0096-1523.23.1.86

Ganis, G. (2013). Visual mental imagery. In *Multisensory imagery* (pp. 9–28). Springer New York. https://doi.org/10.1007/978-1-4614-5879-1_2

Ganis, G., & Kievit, R. (2015). A New Set of Three-Dimensional Shapes for Investigating Mental Rotation Processes: Validation Data and Stimulus Set. *Journal of Open Psychology Data*, *3*. https://doi.org/10.5334/jopd.ai

Ganis, G., Thompson, W., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research*, *20*(2), 226–241. https://doi.org/10.1016/j.cogbrainres.2004.02.012

Gau, S. S. F., & Chiang, H. L. (2013). Association between early attention-deficit/hyperactivity symptoms and current verbal and visuo-spatial short-term memory. *Research in Developmental Disabilities*, *34*(1), 710–720. https://doi.org/10.1016/j.ridd.2012.10.005

Gau, S. S. F., Chiu, C. De, Shang, C. Y., Cheng, A. T. A., & Soong, W. T.

(2009). Executive function in adolescence among children with attention-deficit/ hyperactivity disorder in Taiwan. *Journal of Developmental and Behavioral Pediatrics*, *30*(6), 525–534. https://doi.org/10.1097/DBP.0b013e3181c21c97

Geer, E. A., Quinn, J. M., & Ganley, C. M. (2019). Relations between spatial skills and math performance in elementary school children: A longitudinal investigation. *Developmental Psychology*, *55*(3), 637–652. https://doi.org/10.1037/dev0000649

Gibbons, J. (1993). Nonparametric Measures of Association. In *Nonparametric Measures of Association*. SAGE Publications, Inc. https://doi.org/10.4135/9781412985291

Gilligan, K. A., Flouri, E., & Farran, E. K. (2017). The contribution of spatial ability to mathematics achievement in middle childhood. *Journal of Experimental Child Psychology*, *163*, 107–125. https://doi.org/10.1016/j.jecp.2017.04.016

Gilligan, K. A., Thomas, M. S. C., & Farran, E. K. (2019). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. *Developmental Science*, *23*(4), e12909. https://doi.org/10.1111/desc.12909

Giofrè, D., Mammarella, I. C., Ronconi, L., & Cornoldi, C. (2013). Visuospatial working memory in intuitive geometry, and in academic achievement in geometry. *Learning and Individual Differences*, *23*(1), 114–122. https://doi.org/10.1016/j.lindif.2012.09.012

Giovannini, L., Jacomuzzi, A. C., Bruno, N., Semenza, C., & Surian, L. (2009). Distance perception in autism and typical development. *Perception*, *38*(3), 429–441. https://doi.org/10.1068/p6266

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. In *Trends in Neurosciences* (Vol. 15, Issue 1, pp. 20–25). Elsevier Current Trends. https://doi.org/10.1016/0166-2236(92)90344-8

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, *61*(2), 215–231. https://doi.org/10.1093/biomet/61.2.215

Gosling, S. G., & Astle, D. E. (2013). Directing spatial attention to locations within remembered and imagined mental representations. *Frontiers in Human Neuroscience*, *7*(APR 2013), 1–8.

https://doi.org/10.3389/fnhum.2013.00154
Gratton, G. (1998). The contralateral organization of visual memory: A theoretical concept and a research tool. *Psychophysiology*, *35*(6), 638–647. https://doi.org/10.1111/1469-8986.3560638
Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10–11), 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6
Guarnera, M., Pellerone, M., Commodari, E., Valenti, G. D., & Buccheri, S. L. (2019). Mental Images and School Learning: A Longitudinal Study on Children. *Frontiers in Psychology*, *10*, 2034. https://doi.org/10.3389/fpsyg.2019.02034
Handy, T. C., Miller, M. B., Schott, B., Shroff, N. M., Janata, P., Van Horn, J. D., Inati, S., Grafton, S. T., & Gazzaniga, M. S. (2004). Visual imagery and memory: Do retrieval strategies affect what the mind's eye sees? *European Journal of Cognitive Psychology*, *16*(5), 631–652. https://doi.org/10.1080/09541440340000457
Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635. https://doi.org/10.1038/nature07832
Hart, H., Radua, J., Nakao, T., Mataix-Cols, D., & Rubia, K. (2013). Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: Exploring task-specific, stimulant medication, and age effects. *JAMA Psychiatry*, *70*(2), 185–198. https://doi.org/10.1001/jamapsychiatry.2013.277
Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736
Hegarty, M., & Kozhevnikov, M. (1999). Types of Visual-Spatial Representations and Mathematical Problem Solving. *Journal of Educational Psychology*, *91*(4), 684–689. https://doi.org/10.1037/0022-0663.91.4.684
Hespos, S. J., & Rochat, P. (1997). Dynamic mental representation in infancy. *Cognition*, *64*(2), 153–188. https://doi.org/10.1016/S0010-0277(97)00029-2
Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads

to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), 9–15. https://doi.org/10.1111/j.1467-7687.2009.00848.x

Hyun, J.-S., & Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychonomic Bulletin & Review*, *14*(1), 154–158. https://doi.org/10.3758/BF03194043

Ishai, A. (2002). Visual Imagery of Famous Faces: Effects of Memory and Attention Revealed by fMRI. *NeuroImage*, *17*(4), 1729–1741. https://doi.org/10.1006/nimg.2002.1330

Ishai, Alumit. (2010). Seeing with the mind's eye : top-down, bottom-up, and conscious awareness. *Biology Reports*, *2*(34), 1–4. https://doi.org/10.3410/B2-34

Ishai, Alumit, Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, *28*(3), 979–990. https://doi.org/10.1016/S0896-6273(00)00168-9

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*, *105*, 61–73. https://doi.org/10.1016/j.cortex.2017.10.014

Jakobson, A., & Kikas, E. (2007). Cognitive functioning in children with and without Attention-deficit/Hyperactivity Disorder with and without comorbid learning disabilities. *Journal of Learning Disabilities*, *40*(3), 194–202. https://doi.org/10.1177/00222194070400030101

Johnson, M. R., & Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. *Frontiers in Human Neuroscience*, *8*(1), 59. https://doi.org/10.3389/fnhum.2014.00059

Johnstone, S. J., Dimoska, A., Smith, J. L., Barry, R. J., Pleffer, C. B., Chiswick, D., & Clarke, A. R. (2007). The development of stop-signal and Go/Nogo response inhibition in children aged 7-12 years: Performance and event-related potential indices. *International Journal of Psychophysiology*, *63*(1), 25–38. https://doi.org/10.1016/j.ijpsycho.2006.07.001

Jongen, E. M. M., Smulders, F. T. Y., & Van Der Heiden, J. S. H. (2007). Lateralized ERP components related to spatial orienting: Discriminating the direction of attention from processing sensory aspects of the cue. *Psychophysiology*, *44*(6), 968–986. https://doi.org/10.1111/j.1469-8986.2007.00557.x

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*(3), 490–501. https://doi.org/10.1037/0033-2909.109.3.490

Kail, R., Pellegrino, J., & Carter, P. (1980). Developmental changes in mental rotation. *Journal of Experimental Child Psychology*, *29*(1), 102–116. https://doi.org/10.1016/0022-0965(80)90094-6

Karalunas, S. L., Hawkey, E., Gustafsson, H., Miller, M., Langhorst, M., Cordova, M., Fair, D., & Nigg, J. T. (2018). Overlapping and Distinct Cognitive Impairments in Attention-Deficit/Hyperactivity and Autism Spectrum Disorder without Intellectual Disability. *Journal of Abnormal Child Psychology*, *46*(8), 1705–1716. https://doi.org/10.1007/s10802-017-0394-2

Karmiloff-Smith, A. (1994). Précis of Beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences*, *17*(4), 693–707. https://doi.org/10.1017/S0140525X00036621

Kasper, L. J., Alderson, R. M., & Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): A meta-analytic review. *Clinical Psychology Review*, *32*, 605–617. https://doi.org/10.1016/j.cpr.2012.07.001

Keogh, R., & Pearson, J. (2011). Mental Imagery and Visual Working Memory. *PLOS ONE*, *6*(12), e29221. https://doi.org/10.1371/journal.pone.0029221

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, *14:7*(12), 1–13. https://doi.org/10.1167/14.12.7.doi

Keogh, R., & Pearson, J. (2017). The blind mind: No sensory visual imagery in aphantasia. *Cortex*. https://doi.org/10.1016/J.CORTEX.2017.10.012

Kinahan, P. E., & Noll, D. C. (1999). A direct comparison between whole-brain PET and BOLD fMRI measurements of single-subject activation response. *NeuroImage*, *9*(4), 430–438. https://doi.org/10.1006/nimg.1998.0420

Klein, I., Paradis, A. L., Poline, J. B., Kosslyn, S. M., & Le Bihan, D. (2000). Transient activity in the human calcarine cortex during visual-mental imagery: An event-related fMRI study. *Journal of Cognitive Neuroscience*, *12*(2), 15–23. https://doi.org/10.1162/089892900564037

Klein, Isabelle, Dubois, J., Mangin, J. F., Kherif, F., Flandin, G., Poline, J. B., Denis, M., Kosslyn, S. M., & Le Bihan, D. (2004). Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, *22*(1), 26–31. https://doi.org/10.1016/j.cogbrainres.2004.07.006

Koenig-Robert, R., & Pearson, J. (2019). Decoding the contents and strength of imagery before volitional engagement. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39813-y

Kofler, M. J., Irwin, L. N., Soto, E. F., Groves, N. B., Harmon, S. L., & Sarver, D. E. (2019). Executive Functioning Heterogeneity in Pediatric ADHD. *Journal of Abnormal Child Psychology*, *47*(2), 273–286. https://doi.org/10.1007/s10802-018-0438-2

Kosslyn, S., Ball, T., & Reiser, B. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 47–60. http://psycnet.apa.org/journals/xhp/4/1/47/

Kosslyn, S. M. (1980). *Image and Mind*. Havard University Press.

Kosslyn, S. M. (1996). *Image and brain: the resolution of the imagery debate*. MIT Press.

Kosslyn, S. M. (2005). Mental images and the brain. In *Cognitive Neuropsychology* (Vol. 22, Issues 3–4, pp. 333–347). Psychology Press. https://doi.org/10.1080/02643290442000130

Kosslyn, S. M., Alpert, N. M., Thompson, W., Maljkovic, V., Weise, S. B., Chabris, C. F., Hamilton, S. E., Rauch, S. L., & Buonanno, F. S. (1993). Visual Mental Imagery Activates Topographically Organized Visual Cortex: PET Investigations. *Journal of Cognitive Neuroscience*, *5*(3), 263–287. https://doi.org/10.1162/jocn.1993.5.3.263

Kosslyn, S. M., Brunn, J., Cave, K. R., & Wallach, R. (1984a). Individual differences in mental imagery ability: A computational analysis. *Cognition*, *18*(1–3), 195–243. https://doi.org/10.1016/0010-0277(84)90025-8

Kosslyn, S. M., Brunn, J., Cave, K. R., & Wallach, R. W. (1984b). Individual differences in mental imagery ability: A computational analysis. *Cognition*, *18*(1–3), 195–243.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2003). Mental imagery: against the nihilistic hypothesis. *Trends in Cognitive Sciences*, *7*(3), 109–111. https://doi.org/10.1016/S1364-6613(03)00021-4

Kosslyn, S. M., Margolis, J. A., Barrett, A. M., Goldknopf, E. J., & Daly, P. F. (1990). Age Differences in Imagery Abilities. *Child Development*, *61*(4), 995–1010. https://doi.org/10.1111/j.1467-8624.1990.tb02837.x

Kosslyn, S. M., Murphy, G. L., Bemesderfer, M. E., & Feinstein, K. J. (1977). Category and Continuum in Mental Comparisons. *Journal of Experimental Psychology: General*, *106*(4), 341–375.

Kosslyn, S. M., Shin, L. M., Thompson, W., McNally, R. J., Rauch, S. L., Pitman, R. K., & Alpert, N. M. (1996). Neural effects of visualizing and perceiving aversive stimuli: A PET investigation. *NeuroReport*, *7*(10), 1569–1576. https://doi.org/10.1097/00001756-199607080-00007

Kosslyn, S. M., Thompson, W., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496–498. https://doi.org/10.1038/378496a0

Kosslyn, S. M., & Thompson, W. L. (2003). When Is Early Visual Cortex Activated During Visual Mental Imagery? *Psychological Bulletin*, *129*(5), 723–746. https://doi.org/10.1037/0033-2909.129.5.723

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford University Press.

Kosslyn, S. M., Thompson, W., Shephard, J. M., Ganis, G., Bell, D., Danovitch, J., Wittenberg, L. A., & Alpert, N. M. (2004). Brain rCBF and performance in visual imagery tasks: Common and distinct processes. *European Journal of Cognitive Psychology*, *16*(5), 696–716. https://doi.org/10.1080/09541440340000475

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. In *Nature Reviews Neuroscience* (Vol. 12, Issue 4, pp. 217–230). NIH Public Access. https://doi.org/10.1038/nrn3008

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. In *Trends in Cognitive Sciences* (Vol. 17, Issue 1, pp. 26–49). NIH Public Access. https://doi.org/10.1016/j.tics.2012.10.011

Lakens, D. (2021). *Sample Size Justification*. https://doi.org/10.31234/OSF.IO/9D3YF

Lambert, S., Sampaio, E., Scheiber, C., & Mauss, Y. (2002). Neural substrates of animal mental imagery: Calcarine sulcus and dorsal pathway involvement - An fMRI study. *Brain Research*, *924*(2), 176–183. https://doi.org/10.1016/S0006-8993(01)03232-2

Le Bihan, D., Turner, R., Zeffiro, T. A., Cuenod, C. A., Jezzard, P., & Bonnerot, V. (1993). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(24), 11802–11805. https://doi.org/10.1073/pnas.90.24.11802

Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, *59*(4), 4064–4073. https://doi.org/10.1016/J.NEUROIMAGE.2011.10.055

Lee, S. H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, *59*(4), 4064–4073. https://doi.org/10.1016/j.neuroimage.2011.10.055

Lesch, K. P. (2018). 'Shine bright like a diamond!': is research on high-functioning ADHD at last entering the mainstream? In *Journal of Child Psychology and Psychiatry and Allied Disciplines* (Vol. 59, Issue 3, pp. 191–192). Blackwell Publishing Ltd. https://doi.org/10.1111/jcpp.12887

Liesefeld, H. R., & Zimmer, H. D. (2013). Think spatial: The representation in mental rotation is nonvisual. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(1), 167–182. https://doi.org/10.1037/a0028904

Linn, M. C., & Petersen, A. C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development*, *56*(6), 1479. https://doi.org/10.2307/1130467

Logie, R. H. (1995). *Visuo-spatial working memory*. L. Erlbaum Associates.

Lorenc, E. S., Lee, T. G., Chen, A. J.-W., & D'Esposito, M. (2015). The Effect of Disruption of Prefrontal Cortical Function with Transcranial Magnetic Stimulation on Visual Working Memory. *Frontiers in Systems Neuroscience*, *9*(DEC), 169. https://doi.org/10.3389/fnsys.2015.00169

Lowrie, T., Logan, T., & Hegarty, M. (2019). The Influence of Spatial Visualization Training on Students' Spatial Reasoning and Mathematics Performance. *Journal of Cognition and Development*, *20*(5), 729–751. https://doi.org/10.1080/15248372.2019.1653298

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. In *Neuroscience and Biobehavioral Reviews* (Vol. 62, pp. 100–108). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2016.01.003

Lütke, N., & Lange-Küttner, C. (2015). Keeping It in Three Dimensions:

Measuring the Development of Mental Rotation in Children with the Rotated Colour Cube Test (RCCT). *International Journal of Developmental Sciences*, *9*(2), 95–114. https://doi.org/10.3233/DEV-14154

Machizawa, M. G., Driver, J., & Watanabe, T. (2020). Gray Matter Volume in Different Cortical Structures Dissociably Relates to Individual Differences in Capacity and Precision of Visual Working Memory. *Cerebral Cortex*, *30*(9), 4759–4770. https://doi.org/10.1093/cercor/bhaa046

Machizawa, M. G., Goh, C. C. W., & Driver, J. (2012). Human Visual Short-Term Memory Precision Can Be Varied at Will When the Number of Retained Items Is Low. *Psychological Science*, *23*(6), 554–559. https://doi.org/10.1177/0956797611431988

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Mammarella, I. C., Pazzaglia, F., & Cornoldi, C. (2008). Evidence for different components in children’s visuospatial working memory. *British Journal of Developmental Psychology*, *26*(3), 337–355. https://doi.org/10.1348/026151007X236061

Mammarella, I. C., Todeschini, M., Englaro, G., Lucangeli, D., & Cornoldi, C. (2012). *Geometria test*. Erickson.

Marks, D. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery*, *19*(3–4), 153–167.

Marmor, G. S. (1975). Development of kinetic images: When does the child first represent movement in mental images? *Cognitive Psychology*, *7*(4), 548–559. https://doi.org/10.1016/0010-0285(75)90022-5

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(4), 377–384.

https://doi.org/10.1097/01.chi.0000153228.72591.73
Martinussen, R., & Tannock, R. (2006). Working memory impairments in children with attention-deficit hyperactivity disorder with and without comorbid language learning disorders. *Journal of Clinical and Experimental Neuropsychology*, *28*(7), 1073–1094. https://doi.org/10.1080/13803390500205700
Massonnié, J., Frasseto, P., Mareschal, D., & Kirkham, N. Z. (2020). Scientific Collaboration with Educators: Practical Insights from an in-Class Noise-Reduction Intervention. *Mind, Brain, and Education*. https://doi.org/10.1111/mbe.12240
Matusz, P. J., & Eimer, M. (2011). Multisensory enhancement of attentional capture in visual search. *Psychonomic Bulletin and Review*, *18*(5), 904–909. https://doi.org/10.3758/s13423-011-0131-8
Matusz, P. J., Wallace, M. T., & Murray, M. M. (2017). A multisensory perspective on object memory. *Neuropsychologia*, *105*(February), 243–252. https://doi.org/10.1016/j.neuropsychologia.2017.04.008
Mazard, A., Laou, L., Joliot, M., & Mellet, E. (2005). Neural impact of the semantic content of visual mental images and visual percepts. *Cognitive Brain Research*, *24*(3), 423–435. https://doi.org/10.1016/j.cogbrainres.2005.02.018
McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological Measures of Maintaining Representations in Visual Working Memory. *Cortex*, *43*(1), 77–94. https://doi.org/10.1016/S0010-9452(08)70447-7
Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom-up meets top-down: Neuronal interactions during perception and imagery. *Cerebral Cortex*, *14*(11), 1256–1265. https://doi.org/10.1093/cercor/bhh087
Mellet, E., Tzourio-Mazoyer, N., Bricogne, S., Mazoyer, B., Kosslyn, S. M., & Denis, M. (2000). Functional anatomy of high-resolution visual mental imagery. *Journal of Cognitive Neuroscience*, *12*(1), 98–109. https://doi.org/10.1162/08989290051137620
Mellet, E., Tzourio, N., Denis, M., & Mazoyer, B. (1995). A positron emission tomography study of visual and mental spatial exploration. *Journal of Cognitive Neuroscience*, *7*(4), 433–445. https://doi.org/10.1162/jocn.1995.7.4.433

Miller, B. T., & D’Esposito, M. (2005). Searching for “the top” in top-down control. *Neuron*, *48*(4), 535–538. https://doi.org/10.1016/j.neuron.2005.11.002

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. https://doi.org/10.1016/0166-2236(83)90190-X

Mix, K. S., Levine, S. C., Cheng, Y. L., Stockton, J. D. S., & Bower, C. (2020). Effects of Spatial Training on Mathematics in First and Sixth Grade Children. *Journal of Educational Psychology*, *113*(2), 304–314. https://doi.org/10.1037/edu0000494

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

Möhring, W., Newcombe, N. S., & Frick, A. (2016). Using Mental Transformation Strategies for Spatial Scaling: Evidence From a Discrimination Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(9), 1473–1479. https://doi.org/10.1037/xlm0000240

Moore, D. S., & Johnson, S. P. (2011). Mental Rotation of Dynamic, Three-Dimensional Stimuli by 3-Month-Old Infants. *Infancy*, *16*(4), 435–445. https://doi.org/10.1111/j.1532-7078.2010.00058.x

Moriya, J. (2018). Visual mental imagery influences attentional guidance in a visual-search task. *Attention, Perception, and Psychophysics*, *80*(5), 1127–1142. https://doi.org/10.3758/s13414-018-1520-0

Morris, S., Farran, E. K., & Dumontheil, I. (2019). Field Independence Associates with Mathematics and Science Performance in 5- to 10-Year-Olds after Accounting for Domain-General Factors. *Mind, Brain, and Education*, *13*(4), 268–278. https://doi.org/10.1111/mbe.12214

Murray, A. L., Booth, T., Ribeaud, D., & Eisner, M. (2018). Disagreeing about development: An analysis of parent-teacher agreement in ADHD symptom trajectories across the elementary school years. *International Journal of Methods in Psychiatric Research*, *27*(3). https://doi.org/10.1002/mpr.1723

Murray, A. M., Nobre, A. C., & Stokes, M. G. (2011). Markers of preparatory attention predict visual short-term memory performance. *Neuropsychologia*, *49*(6), 1458–1465. https://doi.org/10.1016/j.neuropsychologia.2011.02.016

Murray, M. M., Michel, C. M., De Peralta, R. G., Ortigue, S., Brunet, D., Andino, S. G., & Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage*, *21*(1), 125–135. https://doi.org/10.1016/j.neuroimage.2003.09.035

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016). The multisensory function of the human primary visual cortex. *Neuropsychologia*, *83*, 161–169. https://doi.org/10.1016/j.neuropsychologia.2015.08.011

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, *21*(6), 449–461. https://doi.org/10.1016/j.tics.2017.03.010

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, *105*, 215–228. https://doi.org/10.1016/J.NEUROIMAGE.2014.10.018

Neuburger, S., Jansen, P., Heil, M., & Quaiser-pohl, C. (2011). Gender differences in pre-adolescents’ mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences*, *50*(8), 1238–1242. https://doi.org/10.1016/j.paid.2011.02.017

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents’ mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences*, *50*(8), 1238–1242. https://doi.org/10.1016/j.paid.2011.02.017

Ngiam, W. X. Q., Adam, K. C. S., Quirk, C., Vogel, E. K., & Awh, E. (2021). Estimating the statistical power to detect set-size effects in contralateral delay activity. *Psychophysiology*, *58*(5), e13791. https://doi.org/10.1111/psyp.13791

Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, *57*(11), 1224–1230. https://doi.org/10.1016/j.biopsych.2004.08.025

Nikolas, M. A., & Nigg, J. T. (2013). Neuropsychological performance and attention-deficit hyperactivity disorder subtypes and symptom dimensions. *Neuropsychology*, *27*(1), 107–120. https://doi.org/10.1037/a0030685

Nikolas, M. A., & Nigg, J. T. (2015). Moderators of Neuropsychological Mechanism in Attention-Deficit Hyperactivity Disorder. *Journal of Abnormal Child Psychology*, *43*(2), 271–281. https://doi.org/10.1007/s10802-014-9904-7

Nobre, A. C., Sebestyen, G. N., & Miniussi, C. (2000). The dynamics of shifting visuospatial attention revealed by event-related potentials. *Neuropsychologia*, *38*(7), 964–974. https://doi.org/10.1016/S0028-3932(00)00015-4

Norman, J. F., Adkins, O. C., & Pedersen, L. E. (2016). The visual perception of distance ratios in physical space. *Vision Research*, *123*, 1–7. https://doi.org/10.1016/j.visres.2016.03.009

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–1023. https://doi.org/10.1162/08989290051137549

Orban, S. A., Rapport, M. D., Friedman, L. M., Eckrich, S. J., & Kofler, M. J. (2018). Inattentive Behavior in Boys with ADHD during Classroom Instruction: the Mediating Role of Working Memory Processes. *Journal of Abnormal Child Psychology*, *46*(4), 713–727. https://doi.org/10.1007/s10802-017-0338-x

Ortega, R., López, V., Carrasco, X., Escobar, M. J., García, A. M., Parra, M. A., & Aboitiz, F. (2020). Neurocognitive mechanisms underlying working memory encoding and retrieval in Attention-Deficit/Hyperactivity Disorder. *Scientific Reports*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-64678-x

Paivio, A. (1978). Comparisons of mental clocks. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 61–71. https://doi.org/10.1037/0096-1523.4.1.61

Patros, C. H. G., Alderson, R. M., Lea, S. E., Tarle, S. J., Kasper, L. J., & Hudec, K. L. (2015). Visuospatial working memory underlies choice-impulsivity in boys with attention-deficit/hyperactivity disorder. *Research in Developmental Disabilities*, *38*, 134–144. https://doi.org/10.1016/j.ridd.2014.12.016

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S. M. A., Heyes, S. B., & Holmes, E. A. (2013). Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework. *Clinical Psychology Review*, *33*(1), 1–23. https://doi.org/10.1016/j.cpr.2012.09.001

Pearson, J., Clifford, C., & Tong, F. (2008). The Functional Impact of Mental Imagery on Conscious Perception. *Current Biology*, *18*(13), 982–986. https://doi.org/10.1016/j.cub.2008.05.048

Pearson, J., & Keogh, R. (2019). Redefining Visual Working Memory: A Cognitive-Strategy, Brain-Region Approach. *Current Directions in Psychological Science*, *28*(3), 266–273. https://doi.org/10.1177/0963721419835210

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *PNAS*, *112*(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

Pearson, J., Rademaker, R., & Tong, F. (2011). Evaluating the Mind's Eye: The Metacognition of Visual Imagery. *Psychological Science*, *22*(12), 1535–1542.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. In *British Medical Journal* (Vol. 316, Issue 7139, pp. 1236–1238). BMJ Publishing Group. https://doi.org/10.1136/bmj.316.7139.1236

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Policardi, E., Perani, D., Zago, S., Grassi, F., Fazio, F., & Làdavas, E. (1996). Failure to evoke visual images in a case of long-lasting cortical blindness. *Neurocase*, *2*(5), 381–394. https://doi.org/10.1080/13554799608402413

Pounder, Z., Jacob, J., Jacobs, C., Loveday, C., Towell, T., & Silvanto, J. (2018). Mental rotation performance in aphantasia. *Journal of Vision*, *18*(10), 1123. https://doi.org/10.1167/18.10.1123

Prime, D. J., & Jolicoeur, P. (2010). Mental rotation requires visual short-term memory: evidence from human electric cortical activity. *Journal of Cognitive Neuroscience*, *22*(11), 2437–2446. https://doi.org/10.1162/jocn.2009.21337

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*(1), 16–45.

https://doi.org/10.1037/0033-295X.88.1.16
Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioural and Brain Sciences*, *25*, 157–237. https://doi.org/10.1017/S0140525X02280048
Ragni, F., Tucciarelli, R., Andersson, P., & Lingnau, A. (2020). Decoding stimulus identity in occipital, parietal and inferotemporal cortices during visual mental imagery. *Cortex*, *127*, 371–387. https://doi.org/10.1016/j.cortex.2020.02.020
Rapport, M. D., Alderson, R. M., & Sims, V. (2008). Working Memory Deficits in Boys with Attention-deficit / Hyperactivity Disorder (ADHD): The Contribution of Central Executive and Subsystem Processes. *Journal of Abnormal Child Psychology*, *36*(6), 825–837. https://doi.org/10.1007/s10802-008-9215-y
Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review*, *33*(8), 1237–1252. https://doi.org/10.1016/j.cpr.2013.08.005
Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales. Section 2: The coloured progressive matrices*. Oxford Psychologists Press.
Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: Decoding category information during mental imagery. *NeuroImage*, *50*(2), 818–825. https://doi.org/10.1016/j.neuroimage.2009.11.084
Reeder, R. R. (2017). Individual differences shape the content of visual representations. *Vision Research*, *141*, 266–281. https://doi.org/10.1016/j.visres.2016.08.008
Riečanský, I., & Jagla, F. (2008). Linking performance with brain potentials: Mental rotation-related negativity revisited. *Neuropsychologia*, *46*(13), 3069–3073. https://doi.org/10.1016/j.neuropsychologia.2008.06.016
Riečanský, I., Tomova, L., Katina, S., Bauer, H., Fischmeister, F. P. S., & Lamm, C. (2013). Visual image retention does not contribute to modulation of event-related potentials by mental rotation. *Brain and Cognition*, *83*(2), 163–170. https://doi.org/10.1016/j.bandc.2013.07.011
Roland, P. E., & Gulyás, B. (1995). Visual memory, visual imagery, and visual recognition of large field patterns by the human brain: Functional anatomy

by positron emission tomography. *Cerebral Cortex*, *5*(1), 79–93. https://doi.org/10.1093/cercor/5.1.79

Rosenbaum, R. S., McKinnon, M. C., Levine, B., & Moscovitch, M. (2004). Visual imagery deficits, impaired strategic retrieval, or memory loss: disentangling the nature of an amnesic person's autobiographical memory deficit. *Neuropsychologia*, *42*(12), 1619–1635. https://doi.org/10.1016/j.neuropsychologia.2004.04.010

Rosser, R. A., Ensing, S. S., & Mazzeo, J. (1985). The role of stimulus salience in young children's ability to discriminate two-dimensional rotations: Reflections on a paradigm. *Contemporary Educational Psychology*, *10*(2), 95–103. https://doi.org/10.1016/0361-476X(85)90010-4

Sabbah, P., Simond, G., Levrier, O., Habib, M., Trabaud, V., Mur Ay Ama, N., Mazoyer, B. M., Briant, J. F., Ravbaud, C. H., & Salomon, G. (1995). Functional magnetic resonance imaging at 1.5 T during sensorimotor and cognitive task. *European Neurology*, *35*(3), 131–136. https://doi.org/10.1159/000117108

Sadoski, M, & Quast, Z. (1990). Reader Response and Long-Term Recall for Journalistic Text: The Roles of Imagery, Affect, and Importance. *Reading Research Quarterly*, *25*(4), 256–272.

Sadoski, Mark, & Paivio, A. (2004). *Dual Coding Theory in Reading*. 1329–1362.

Sala, G., & Gobet, F. (2017). Working memory training in typically developing children: A multilevel meta-analysis. *Psychonomic Bulletin and Review*, *53*(4), 671–685. https://doi.org/10.3758/s13423-019-01681-y

Salin, P. A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. In *Physiological Reviews* (Vol. 75, Issue 1, pp. 107–154). American Physiological Society. https://doi.org/10.1152/physrev.1995.75.1.107

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67. https://doi.org/10.1016/j.visres.2016.09.010

Shepard, R. N., & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science*, *171*, 701–703.

Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting Attention Within Visual Short-Term Memory: Development and Mechanisms. *Child*

*Development*, *85*(2), 578–592. https://doi.org/10.1111/cdev.12150

Shin, L. M., McNally, R. J., Kosslyn, S. M., Thompson, W. L., Rauch, S. L., Alpert, N. M., Metzger, L. J., Lasko, N. B., Orr, S. P., & Pitman, R. K. (1999). Regional cerebral blood flow during script-driven imagery in childhood sexual abuse-related PTSD: A PET investigation. *American Journal of Psychiatry*, *156*(4), 575–584. https://doi.org/10.1176/ajp.156.4.575

Silk, T., Vance, A., Rinehart, N., Egan, G., O'Boyle, M., Bradshaw, J. L., & Cunnington, R. (2005). Fronto-parietal activation in attention-deficit hyperactivity disorder, combined type: Functional magnetic resonance imaging study. *British Journal of Psychiatry*, *187*(SEPT.), 282–283. https://doi.org/10.1192/bjp.187.3.282

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007

Simmons, F. R., Willis, C., & Adams, A. M. (2012). Different components of working memory have different relationships with different mathematical skills. *Journal of Experimental Child Psychology*, *111*(2), 139–155. https://doi.org/10.1016/j.jecp.2011.08.011

Simone, A. N., Bédard, A. C. V., Marks, D. J., & Halperin, J. M. (2015). Good holders, bad shufflers: An examination of working memory processes and modalities in children with and without attention-deficit/hyperactivity disorder. *Journal of the International Neuropsychological Society*, *22*(1), 1–11. https://doi.org/10.1017/S1355617715001010

Simone, A. N., Marks, D. J., Bédard, A. C., & Halperin, J. M. (2018). Low Working Memory rather than ADHD Symptoms Predicts Poor Academic Achievement in School-Aged Children. *Journal of Abnormal Child Psychology*, *46*(2), 277–290. https://doi.org/10.1007/s10802-017-0288-3

Sjöwall, D., Bohlin, G., Rydell, A. M., & Thorell, L. B. (2017). Neuropsychological deficits in preschool as predictors of ADHD symptoms and academic achievement in late adolescence. *Child Neuropsychology*. https://doi.org/10.1080/09297049.2015.1063595

Sjöwall, D., Roth, L., Lindqvist, S., & Thorell, L. B. (2013). Multiple deficits in ADHD: Executive dysfunction, delay aversion, reaction time variability, and emotional deficits. *Journal of Child Psychology and Psychiatry and Allied*

*Disciplines*, *54*(6), 619–627. https://doi.org/10.1111/jcpp.12006

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex*, *15*(10), 1570–1583. https://doi.org/10.1093/cercor/bhi035

Spagna, A., Hajhajate, D., Liu, J., & Bartolomeo, P. (2021). Visual mental imagery engages the left fusiform gyrus, but not the early visual cortex: a meta-analysis of neuroimaging evidence. *Neuroscience & Biobehavioral Reviews*, *122*, 201–217. https://doi.org/10.1016/j.neubiorev.2020.12.029

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89. https://doi.org/10.1016/j.tics.2013.12.001

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, *29*(5), 1565–1572. https://doi.org/10.1523/JNEUROSCI.4657-08.2009

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5*(11), 1024–1037. https://doi.org/10.1167/5.11.8

Swanson, H. L. (2015). Cognitive strategy interventions improve word problem solving and working memory in children with math disabilities. *Frontiers in Psychology*, *6*, 1099. https://doi.org/10.3389/fpsyg.2015.01099

Thelen, A., Matusz, P. J., & Murray, M. M. (2014). Multisensory context portends object memory. *Current Biology*, *24*(16), 734–735. https://doi.org/10.1016/j.cub.2014.06.040

Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J. B., Lebihan, D., & Dehaene, S. (2006). Inverse retinotopy: Inferring the visual content of images from brain activation patterns. *NeuroImage*, *33*(4), 1104–1116. https://doi.org/10.1016/j.neuroimage.2006.06.062

Thompson, W. L., Hsiao, Y., & Kosslyn, S. M. (2011). Dissociation between visual attention and visual mental imagery. *Journal of Cognitive Psychology*, *23*(2), 256–263. https://doi.org/10.1080/20445911.2011.477810

Thompson, W. L., Kosslyn, S. M., Sukel, K. E., & Alpert, N. M. (2001). Mental imagery of high- and low-resolution gratings activates Area 17. *NeuroImage*, *14*(2), 454–464. https://doi.org/10.1006/nimg.2001.0803

Thurley, K., & Schild, U. (2018). Time and distance estimation in children using an egocentric navigation task. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-36234-1

Tillman, C., Eninger, L., Forssman, L., & Bohlin, G. (2011). The relation between working memory components and ADHD symptoms from a developmental perspective. *Developmental Neuropsychology*, *36*(2), 181–198. https://doi.org/10.1080/87565641.2010.549981

Tong, F. (2013). Imagery and visual working memory: One and the same? *Trends in Cognitive Sciences*, *17*(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

Trojano, L. (2000). Matching Two Imagined Clocks: the Functional Anatomy of Spatial Analysis in the Absence of Visual Stimulation. *Cerebral Cortex*, *10*(5), 473–481. https://doi.org/10.1093/cercor/10.5.473

Ungerleider, L. G., & Bell, A. H. (2011). Uncovering the visual “alphabet”: Advances in our understanding of object perception. *Vision Research*, *51*(7), 782–799. https://doi.org/10.1016/j.visres.2010.10.002

Ungerleider, L. G., & Haxby, J. V. (1994). “What” and “where” in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157–165. https://doi.org/10.1016/0959-4388(94)90066-3

Vaidya, C. J., You, X., Mostofsky, S., Pereira, F., Berl, M. M., & Kenworthy, L. (2020). Data-driven identification of subtypes of executive function across typical development, attention deficit hyperactivity disorder, and autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *61*(1), 51–61. https://doi.org/10.1111/jcpp.13114

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science*, *255*(5043), 419–423. https://doi.org/10.1126/science.1734518

Vance, A., Silk, T. J., Casey, M., Rinehart, N. J., Bradshaw, J. L., Bellgrove, M. A., & Cunnington, R. (2007). Right parietal dysfunction in children with attention deficit hyperactivity disorder, combined type: A functional MRI study. *Molecular Psychiatry*, *12*(9), 826–832. https://doi.org/10.1038/sj.mp.4001999

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. https://doi.org/10.1038/nature02447

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/ hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2005.02.006

Williams, G. A., & Kibwoski, F. (2016). Latent Class Analysis and Latent Profile Analysis. In *Handbook of Methodological Approaches to Community-Based Research: Qualitative, Quantitative, and Mixed Methods* (pp. 143–151). Oxford University Press.

Williams, J., Omizzolo, C., Galea, M. P., & Vance, A. (2013). Motor imagery skills of children with Attention Deficit Hyperactivity Disorder and Developmental Coordination Disorder. *Human Movement Science*, *32*(1), 121–135. https://doi.org/10.1016/j.humov.2012.08.003

Wimmer, M. C., Maras, K. L., Robinson, E. J., Doherty, M. J., & Pugeault, N. (2015). How visuo-spatial mental imagery develops: Image generation and maintenance. *PLoS ONE*, *10*(11), 1–20. https://doi.org/10.1371/journal.pone.0142566

Wimmer, M. C., Maras, K. L., Robinson, E. J., & Thomas, C. (2016). The format of children’s mental images: Evidence from mental scanning. *Cognition*, *154*, 49–54. https://doi.org/10.1016/j.cognition.2016.05.017

Wimmer, M. C., Robinson, E. J., & Doherty, M. J. (2017). Are developments in mental scanning and mental rotation related? *PLoS ONE*, *12*(2), 1–12. https://doi.org/10.1371/journal.pone.0171762

Winlove, C. I. P., Milton, F., Ranson, J., Fulford, J., Mackisack, M., Macpherson, F., & Zeman, A. (2018). The neural correlates of visual imagery: A co-ordinate-based meta-analysis. *Cortex*, *105*, 4–25. https://doi.org/10.1016/j.cortex.2017.12.014

Wojciulik, E., & Kanwisher, N. (1999). The Generality of Parietal Involvement in Visual Attention. *Neuron*, *23*(4), 747–764. https://doi.org/10.1016/S0896-6273(01)80033-7

Zago, S., Corti, S., Bersano, A., Baron, P., Conti, G., Ballabio, E., Lanfranconi, S., Cinnante, C., Costa, A., Cappellari, A., & Bresolin, N. (2010). A Cortically Blind Patient With Preserved Visual Imagery. *Cognitive and Behavioral Neurology*, *23*(1), 44–48. https://doi.org/10.1097/WNN.0b013e3181bf2e6e

Zanto, T. P., & Gazzaley, A. (2013). Fronto-parietal network: Flexible hub of cognitive control. *Trends in Cognitive Sciences*, *17*(12), 602–603. https://doi.org/10.1016/j.tics.2013.10.001

Zelazo, P. D., & Carlson, S. M. (2012). Hot and Cool Executive Function in Childhood and Adolescence: Development and Plasticity. *Child Development Perspectives*, *6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery - Congenital aphantasia. *Cortex*, *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

Zeman, A. Z. J., Della Sala, S., Torrens, L. A., Gountouna, V. E., McGonigle, D. J., & Logie, R. H. (2010). Loss of imagery phenomenology with intact visuo-spatial task performance: A case of “blind imagination.” *Neuropsychologia*, *48*(1), 145–155. https://doi.org/10.1016/j.neuropsychologia.2009.08.024

# Appendix

## A.1. Chapter 2

### A.1.1. Sensitivity power analysis

A sensitivity power analysis was conducted using G*Power 3.1 (Faul et al., 2007). with the following parameters: two tails, alpha level = .002, power = 0.8, correlation H0 = 0. The sensitivity curve plot is displayed in Figure A.1.

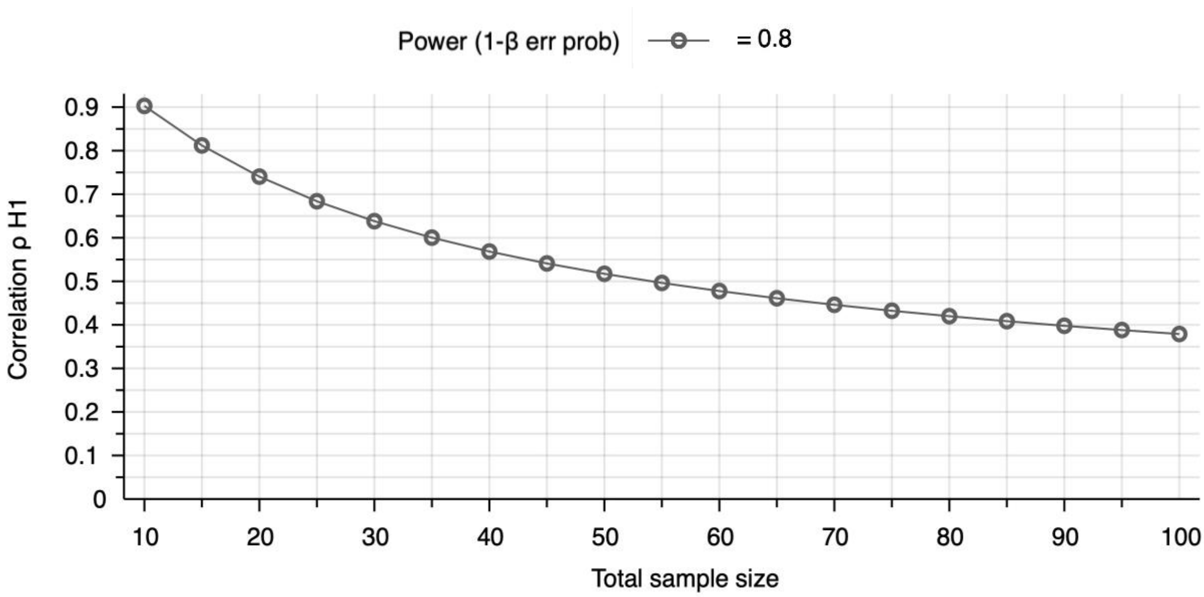



*Figure A.1.1.1:* Sensitivity curve plot of power (0.8) for desired alpha level (.002)

## A.1.2. Tables of response times in milliseconds

Table A.1.2.1.

*Means and standard deviations for response times (RTs) in milliseconds (ms) for image generation and image maintenance*

| | RT measures | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Image Generation High precision | | Image Generation Low precision | | Image Maintenance High precision | | Image Maintenance Low precision | |
| Age group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 9844 | 2840 | 11682 | 3535 | 10670 | 2482 | 13656 | 5680 |
| 8-9 years | 8345 | 2066 | 11553 | 4400 | 8498 | 1997 | 10726 | 2481 |
| 10-11 years | 9212 | 2902 | 11810 | 3977 | 8470 | 2056 | 9927 | 3154 |
| Adults | 8616 | 2733 | 11278 | 4413 | 6322 | 1879 | 7725 | 2475 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

Table A.1.2.2.

*Means and standard deviations for RTs in ms for mental rotation*

| | Degrees of rotation | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | | 45 | | 90 | | 135 | | 180 | |
| Age group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 3622 | 1295 | 4208 | 1461 | 5422 | 3284 | 5440 | 1229 | 5149 | 2082 |
| 8-9 years | 3496 | 1500 | 3803 | 1884 | 4718 | 2777 | 5404 | 2094 | 5904 | 2966 |
| 10-11 years | 2287 | 733 | 2745 | 936 | 3438 | 1642 | 3926 | 1600 | 4354 | 1835 |
| Adults | 1595 | 516 | 1807 | 641 | 2073 | 682 | 2535 | 870 | 2770 | 1174 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

Table A.1.2.3.

*Means and standard deviations for RTs in ms for image scanning*

| | Distances | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 70mm | | 80mm | | 100mm | | 154mm | | 262mm | |
| Age group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 9678 | 7427 | 11223 | 10459 | 10352 | 9085 | 11296 | 7083 | 18064 | 18002 |
| 8-9 years | 9810 | 6102 | 8575 | 3433 | 11176 | 7045 | 11819 | 6417 | 14183 | 6910 |
| 10-11 years | 11602 | 8272 | 9148 | 3112 | 10885 | 5573 | 15398 | 8407 | 17585 | 9035 |
| Adults | 7576 | 3370 | 7163 | 3054 | 7563 | 3638 | 10764 | 4558 | 14776 | 6109 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

Table A.1.2.4.

*Means and standard deviations for RTs in ms for perception control trials*

| | Perception control trial distances | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 70mm | | 81mm | | 100mm | | 154mm | | 262mm | |
| Age group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 6808 | 4340 | 7743 | 5729 | 7219 | 4390 | 9509 | 6404 | 15308 | 7421 |
| 8-9 years | 8335 | 3435 | 9199 | 3982 | 9101 | 5508 | 12455 | 5397 | 16680 | 7253 |
| 10-11 years | 8518 | 4444 | 8331 | 4231 | 9182 | 4684 | 14979 | 8440 | 16903 | 8835 |
| Adults | 7411 | 3426 | 7488 | 3668 | 8547 | 3962 | 12186 | 5472 | 15624 | 7154 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

### A.1.3. Perception control trial analyses

The equivalent analyses were conducted on perception control trials to investigate how accurately individuals perceive distance in this task. The relationship between the actual distance ratios and mean perceived distance ratios (both calculated in the same manner as in image scanning trials) are displayed graphically in Figure A.1.2.1. As in image scanning, visual inspection shows that the ratios are highly, linearly related. Thus, as actual distance increases, perceived distance increases.

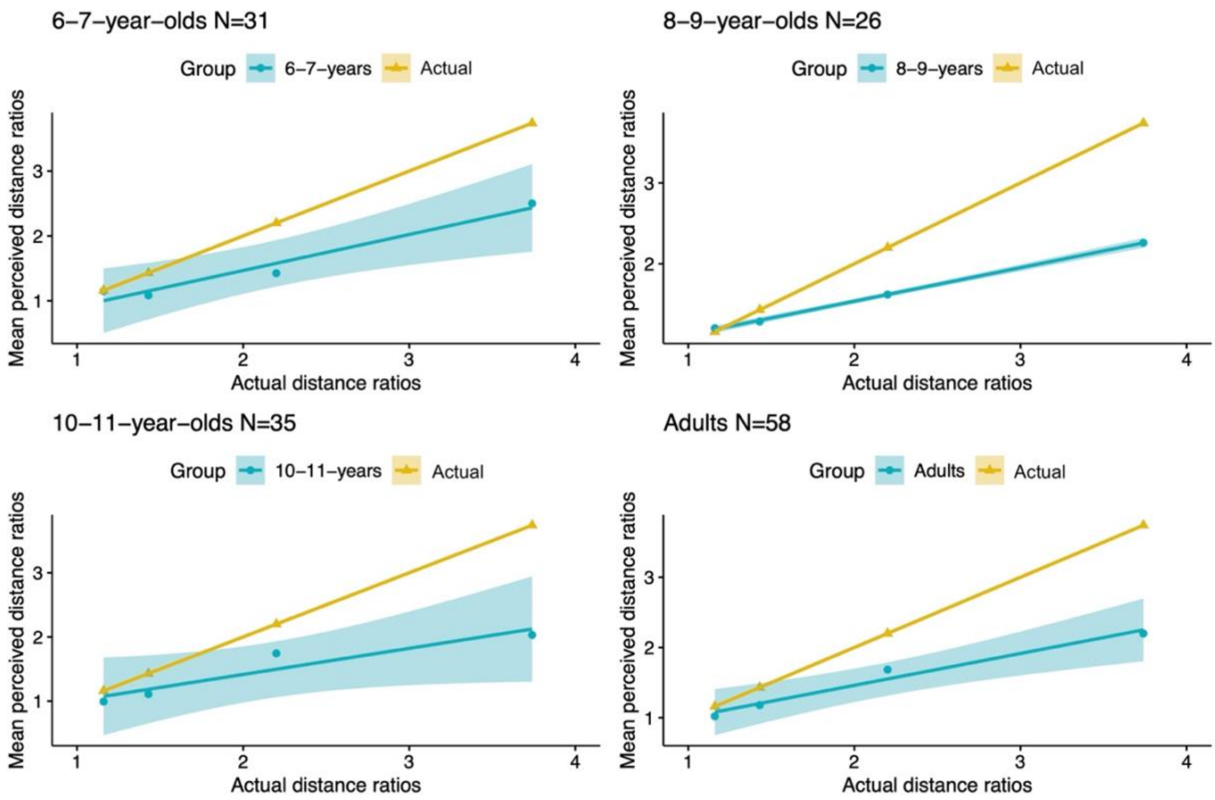



*Figure A.1.3.1:* Scatter plots depicting the relationship between actual distance ratios and mean perceived distance ratios with confidence intervals and best fitting regression lines. Actual distance ratio line added for reference

As detailed in Table A.1.3.1., one sample t tests between the actual distance ratio and perceived distance ratio indicated the exact relationships found in the image scanning analyses: perceived distance ratios were significantly lower than actual distance ratios at all distances in all age groups, except between the 1st perceived ratio and 1st actual ratio in age 6- to 7-year-olds and 8- to 9-year-olds, as well as between the 2nd perceived ratio and 2nd actual ratio in age 8- to 9-year-olds. Thus, suggesting the pattern of underestimating distance seen in image scanning is also seen in individuals' perception of distance in this task. To examine this statistically, a mixed ANOVA was conducted with the dependent variable of mean RT ratio, within-subject factor of ratio type (imaged, perceived) and between-subject factor of age group. This revealed a significant main effect of ratio type ($F(1,146) = 13.73$, $p < .001$, $\eta_p^2 = .09$) suggested lower RT ratios in imaged distance compared to perceived distance. There was no main effect of age group and no interaction between ratio type and age groups (*F*s < 1).

Table A.1.3.1.

*One sample t tests between actual distance and perceived distance*

| Perceived distances | 6-7 years | 8-9 years | 10-11 years | Adults |
|---|---|---|---|---|
| | | *t*(df), *p*, cohens *d* | | |
| Difference from actual ratio 1.16 | -0.11(30), ns, .02 | 0.29(25), ns, .06 | -3.56(34), **, .60 | 5.15(57), ***, .68 |
| Difference from actual ratio 1.43 | -5.53(30), ***, .99 | -0.50(25), ns, .09 | -7.49(34), ***, 1.27 | -5.76(57), ***, .76 |
| Difference from actual ratio 2.2 | -9.83(30), ***, 1.77 | -3.59(25), **, .71 | -6.23(34), ***, 1.05 | -10.16(57), ***, 1.33 |
| Difference from actual ratio 3.74 | -6.86(30), ***, 1.23 | -5.49(25), ***, 1.08 | -14.99(34), ***, 2.53 | -18.30(57), ***, 2.40 |

### A.1.4. Age group comparisons for VWM and attention control

A mixed ANOVA was conducted on sequence length with a within-subject factor of task (maintenance, manipulation) and a between-subject factor of age group. There was a significant main effect of task ($F(1,146) = 5.23$, $p = .024$, $\eta_p^2 = .04$), which showed greater sequence length in maintenance compared to manipulation. There was also a significant main effect of age group ($F(3,146) = 39.64$, $p < .001$, $\eta_p^2 = .45$). Post hoc comparisons showed significantly better performance between all older age groups and younger age groups (*ps* < .05), except for between 8- to 9-year-olds and 6- to 7-year-olds ($p = .308$).

A one-way ANOVA was conducted on Go/No-Go correct mean RT with a between-subject factor of age group. There was significant main effect of RT ($F(3,145) = 58.144$, $p < .001$, $\eta_p^2 = .55$). Post hoc comparisons indicated faster RTs in all older age groups compared to younger age groups (*p*s < .05), except for between 8- to 9-year-olds and 6- to 7-year-olds ($p = .132$). Descriptive statistics for both VWM and Go/No-Go (attention control) are reported in **Table 2.3., Chapter 2, Section 2.3.6.**

# A.2. Chapter 3

## A.2.1. Tables of response times in milliseconds

Table A.2.1.1.

*Means and standard deviations for RTs in ms for mental rotation*

| | Degrees | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | | 45 | | 90 | | 135 | | 180 | |
| Group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 3652 | 1344 | 5553 | 1237 | 4894 | 1586 | 4194 | 1556 | 5034 | 2426 |
| 8-9 years | 3022 | 818 | 5555 | 1623 | 5614 | 1558 | 3381 | 897 | 3946 | 995 |
| 10-11 years | 2230 | 764 | 3706 | 1400 | 4237 | 1546 | 2493 | 920 | 3338 | 1833 |
| ADHD | 2132 | 579 | 3604 | 1410 | 4327 | 1913 | 2488 | 938 | 2801 | 1037 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

Table A.2.1.2.

*Means and standard deviations for RTs in ms for image scanning*

| | Distances | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 70mm | | 80mm | | 100mm | | 154mm | | 262mm | |
| Group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 9823 | 7773 | 10354 | 9526 | 10932 | 7278 | 16836 | 14933 | 11016 | 10342 |
| 8-9 years | 10854 | 6723 | 11852 | 8023 | 12702 | 7049 | 14856 | 6721 | 8852 | 3474 |
| 10-11 years | 12769 | 9539 | 11944 | 6161 | 16960 | 8914 | 19795 | 9952 | 9419 | 3122 |
| ADHD | 12436 | 10581 | 11746 | 11092 | 16047 | 12279 | 17177 | 9520 | 10045 | 7510 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

Table A.2.1.3.

*Means and standard deviations for RTs in ms for perception control trials*

| | Distances | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 70mm | | 81mm | | 100mm | | 154mm | | 262mm | |
| Group | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 6-7 years | 7000 | 4595 | 7215 | 4633 | 9793 | 6766 | 14847 | 7332 | 7534 | 6016 |
| 8-9 years | 7817 | 3373 | 9106 | 6017 | 12721 | 5233 | 14124 | 5957 | 8862 | 4428 |
| 10-11 years | 8631 | 3626 | 9439 | 3874 | 15742 | 7408 | 17400 | 8468 | 8422 | 3822 |
| ADHD | 8962 | 8095 | 10018 | 8599 | 14080 | 10744 | 17238 | 10106 | 9246 | 8751 |

*Note. M* and *SD* represent mean and standard deviation, respectively.

### A.2.2. Perception control trial analyses

One sample t tests between actual ratios and perceived distance ratios showed the same relationships as the imaged ratios in the ADHD group, with the exception of no significant difference between the first actual distance ratio and first imaged distance ratio (ratio 1: $t(18) = -1.34$, $p = .198$, $d = .31$; ratio 2: $t(18) = -2.75$, $p = .013$, $d = .63$; ratio 3: $t(18) = -3.13$, $p = .006$, $d = .72$; ratio 4: $t(18) = -2.48$, $p = .023$, $d = .57$). A mixed ANOVA was conducted with a within-subject factor of perceived distance difference scores (actual distance ratios subtracted from imaged distance ratios) and between-subject factor of group. There was no main effect of group ($F < 1$) and no significant interaction between group and difference score ($F(4.79,133.94) = 1.92$, $p = .099$, $\eta_p^2 = .06$). The significant main effect of difference score was best explained by a significant linear contrast ($F(1,84) = 173.09$, $p < .001$, $\eta_p^2 = .67$) whereby pairwise comparisons revealed significant increases between each difference score (*p*s < .05). This suggests that as demonstrated in **A.1.** in the broader sample, the pattern of underestimating distance in image scanning trials is also observed in individual’s perception of distance in this task (Figure A.2.2.1).

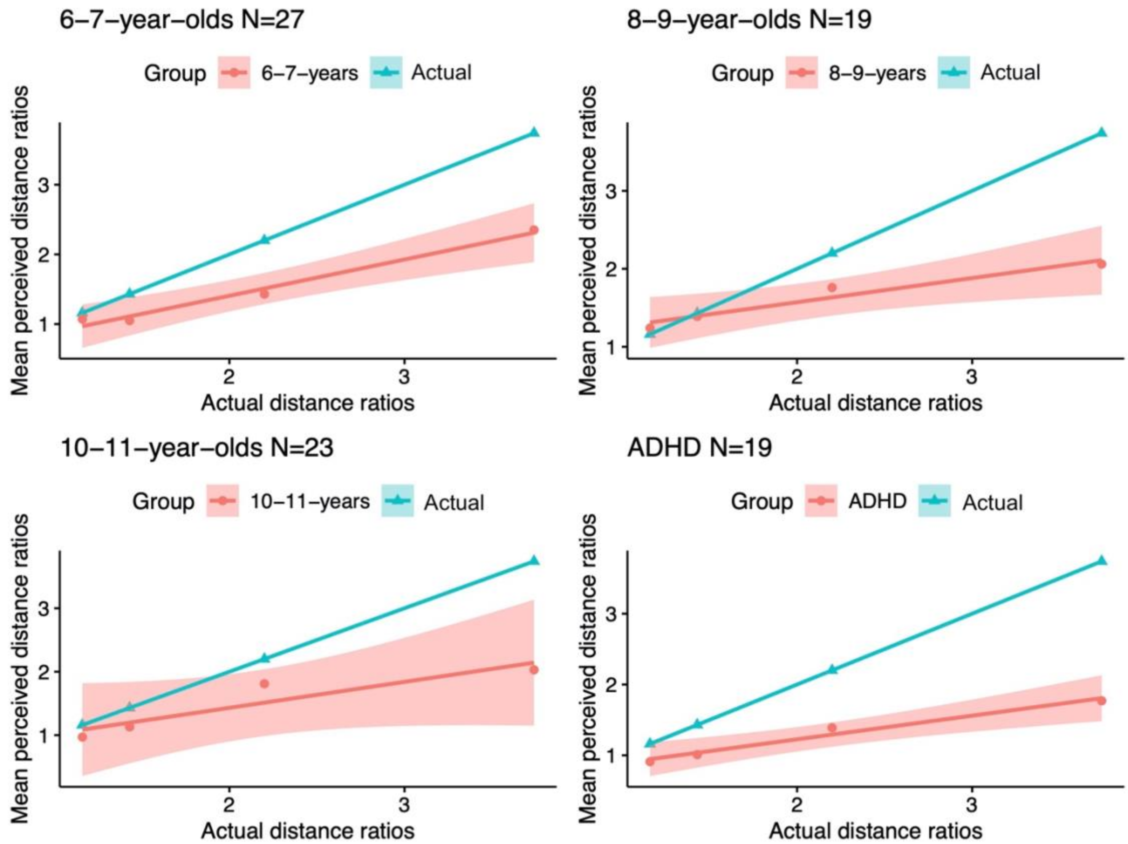



*Figure A.2.2.1:* Scatterplots depicting the relationship between actual distance ratios and mean perceived distance ratios with confidence intervals and best fitting regression lines for each group. Actual distance ratios also plotted for reference

## A.3. Chapter 4

Table A.3.1.

*Descriptive statistics of number of trials in per condition in 1 item trials for the CDA analyses*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean (SD) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) |
| Min. | 5 | 5 | 6 | 5 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 3 | 5 | 6 | 5 | 4 |
| Max. | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

*Note. Q = Quantity, V = Vividness*

Table A.3.2.

*Descriptive statistics of number of trials in per condition in 2 item trials for the CDA analyses*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean (SD) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) |
| Min. | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 3 | 5 | 5 | 3 | 4 | 5 | 4 | 5 | 4 |
| Max. | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

*Note. Q = Quantity, V = Vividness*

Table A.3.3.

*Descriptive statistics of number of trials in per condition in 4 item trials for the CDA analyses*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Coarse | | | | Fine | | | | Coarse | | | | Fine | | | |
| Instruction | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | | Capacity-focused | | Quality-focused | |
| Rating | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V | Q | V |
| Mean (SD) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (1) | 7 (2) |
| Min. | 4 | 4 | 5 | 3 | 3 | 5 | 6 | 4 | 4 | 2 | 3 | 5 | 4 | 5 | 3 | 4 |
| Max. | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

*Note.* Q = Quantity, V = Vividness

Table A.3.4.

*Descriptive statistics of number of trials reported as high and low vividness and non-divergent and divergent quantity ratings included in the ADAN analyses*

| Attended side | Left | | | | Right | | | |
|---|---|---|---|---|---|---|---|---|
| Rating | High vividness | Low vividness | Non-divergent | Divergent | High vividness | Low vividness | Non-divergent | Divergent |
| Mean (SD) | 46 (23) | 34 (22) | 57 (21) | 25 (16) | 48 (24) | 34 (24) | 58 (18) | 24 (15) |
| Min. | 14 | 1 | 21 | 1 | 16 | 3 | 32 | 0 |
| Max. | 82 | 65 | 89 | 53 | 78 | 71 | 92 | 48 |

Table A.3.5.

*Descriptive statistics of number of trials reported as high and low vividness and non-divergent and divergent quantity ratings for 1 item trials in the CDA analyses*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Quality-focused | | | | Capacity-focused | | | | Quality-focused | | | | Capacity-focused | | | |
| Rating | HV | LV | N-D | D | HV | LV | N-D | D | HV | LV | N-D | D | HV | LV | N-D | D |
| Mean (SD) | 8 (4) | 5 (5) | 13 (4) | 2 (4) | 9 (5) | 4 (5) | 12 (3) | 2 (3) | 9 (5) | 4 (6) | 13 (3) | 2 (3) | 10 (5) | 4 (5) | 12 (4) | 2 (3) |
| Min. | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| Max. | 16 | 16 | 16 | 14 | 15 | 15 | 16 | 13 | 16 | 15 | 16 | 11 | 15 | 14 | 16 | 13 |

*Note.* HV = high vividness, LV = low vividness, N-D = non-divergent, D = divergent

Table A.3.6.

*Descriptive statistics of number of trials reported as high and low vividness and non-divergent and divergent quantity ratings for 2 item trials in the CDA analyses*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Quality-focused | | | | Capacity-focused | | | | Quality-focused | | | | Capacity-focused | | | |
| Rating | HV | LV | N-D | D | HV | LV | N-D | D | HV | LV | N-D | D | HV | LV | N-D | D |
| Mean (SD) | 8 (5) | 6 (6) | 12 (3) | 2 (3) | 9 (5) | 5 (5) | 12 (3) | 2 (3) | 7 (6) | 6 (6) | 12 (4) | 2 (2) | 8 (5) | 5 (5) | 12 (4) | 2 (4) |
| Min. | 0 | 0 | 2 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 4 | 0 |
| Max. | 14 | 16 | 16 | 10 | 15 | 16 | 15 | 9 | 15 | 16 | 16 | 7 | 15 | 13 | 16 | 10 |

*Note.* HV = high vividness, LV = low vividness, N-D = non-divergent, D = divergent

Table A.3.7.

*Descriptive statistics of number of trials reported as high and low vividness and non-divergent and divergent quantity ratings for 4 item trials in the CDA analyses*

| Attended side | Left | | | | | | | | Right | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Precision | Quality-focused | | | | Capacity-focused | | | | Quality-focused | | | | Capacity-focused | | | |
| Rating | HV | LV | N-D | D | HV | LV | N-D | D | HV | LV | N-D | D | HV | LV | N-D | D |
| Mean (SD) | 5 (5) | 7 (5) | 5 (5) | 9 (6) | 6 (5) | 8 (5) | 4 (5) | 9 (5) | 5 (5) | 8 (5) | 5 (5) | 9 (5) | 7 (5) | 6 (5) | 4 (5) | 9 (5) |
| Min. | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Max. | 16 | 15 | 15 | 16 | 14 | 15 | 15 | 15 | 16 | 15 | 15 | 15 | 16 | 14 | 15 | 15 |

*Note*. HV = high vividness, LV = low vividness, N-DQ = non-divergent quantity, DQ = divergent quantity